ECONSTOR
Make Your Publications Visible.

A Service of

ZBW Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Mueller-Langer, Frank; Andreoli-Versbach, Patrick

Working Paper

Open Access to Research Data: Strategic Delay and the Ambiguous Welfare Effects of Mandatory Data Disclosure

Munich Discussion Paper, No. 2014-32

Provided in Cooperation with:
University of Munich, Department of Economics

Suggested Citation: Mueller-Langer, Frank; Andreoli-Versbach, Patrick (2014) : Open Access to Research Data: Strategic Delay and the Ambiguous Welfare Effects of Mandatory Data Disclosure, Munich Discussion Paper, No. 2014-32, Ludwig-Maximilians-Universität München, Volkswirtschaftliche Fakultät, München, https://doi.org/10.5282/ubm/epub.21037

This Version is available at:
https://hdl.handle.net/10419/104444

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

WWW.ECONSTOR.EU

Mitglied der Leibniz-Gemeinschaft

LMU LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN
VOLKSWIRTSCHAFTLICHE FAKULTÄT

Frank Mueller-Langer und Patrick Andreoli-Versbach:

Open Access to Research Data: Strategic Delay and the Ambiguous Welfare Effects of Mandatory Data Disclosure

Munich Discussion Paper No. 2014-32

Department of Economics
University of Munich

Volkswirtschaftliche Fakultät
Ludwig-Maximilians-Universität München

Online at http://epub.ub.uni-muenchen.de/21037/

Open Access to Research Data: Strategic Delay and the Ambiguous Welfare Effects of Mandatory Data Disclosure

Frank Mueller-Langer[a], Patrick Andreoli-Versbach[a,b]

[a]Max Planck Institute for Innovation and Competition

[b]Ludwig Maximilian University of Munich, Department of Economics

20 June 2014

Abstract

Mandatory data disclosure is an essential feature for credible empirical work but comes at a cost: First, authors might invest less in data generation if they are not the full residual claimants of their data after their first publication. Second, authors might "strategically delay" the time of submission of papers in order to fully exploit their data in subsequent research. We analyze a three-stage model of publication and data disclosure. We derive exact conditions for positive welfare effects of mandatory data disclosure. However, we find that the transition to mandatory data disclosure has negative welfare properties if authors delay strategically.

Keywords: Data disclosure policy, strategic delay, welfare effects

JEL classification: B40, C80, L59

1 Introduction

> "It is the policy of the American Economic Review to publish papers only if the data used in the analysis are clearly and precisely documented and are readily available to any researcher for purposes of replication. Authors of accepted papers that contain empirical work, simulations, or experimental work must provide to the Review, prior to publication, the data, programs, and other details of the computations sufficient to permit replication. These will be posted on the AER Web site."[1]

Data sharing is an essential feature for the scientific principle of credibility, replication and further research. It guarantees that research methods used to produce the results are known and that incorrect results can be withdrawn from the cumulative body of knowledge (Anderson et al., 2008; Dewald et al., 1986; Vlaeminck and Wagner, 2014). Also, replicable research fosters learning and thus facilitates the development of subsequent research, which boosts scientific advancement. However, whereas a large majority of researchers seems to recognize the importance of data sharing, they are reluctant to apply this principle in practice (Nelson, 2009). Recently, some of the major economics journals have introduced mandatory data disclosure policies which require authors to share their data prior to publication. While these policies provide a strong institutional response in order to overcome low data sharing rates they may have detrimental effects for researchers. Authors who invest in costly data generation, e.g., collecting data and programming, are not the full residual claimants of the data after the first publication and might decrease their initial effort to generate it. In addition, they may strategically delay the submission of papers in order to fully exploit their data. In contrast, disclosed data might increase the value of a publication for its author by increasing its credibility. It may also generate positive effects for the scientific community as the data could be used for subsequent research and validation.

In this paper we first derive the optimal effort choices to generate data under different data policies. We then compare the optimal effort choices and publication strategies (no delay versus strategic delay) across different policies in order to derive welfare implications. We set up a three-stage model of publication and data disclosure to analyze the interaction between a data-creating researcher and a competing researcher and study the incentive and welfare effects of data disclosure. The former researcher exerts effort to create the data in the initial stage. She may then either publish her paper(s) sequentially in the next two stages

[1] http://www.aeaweb.org/aer/data.php (last accessed 6 June 2014). Among others, the Journal of Economic Perspectives, Review of Economic Studies, Journal of Political Economy, and Econometrica adopt virtually the same policy.

or altogether in the last stage to avoid self-induced competition associated with disclosed data. Our motivating example is a researcher who exerts effort to create a unique (possibly hand-collected) data set in order to pursue a novel research agenda, i.e., a set of papers using the same data. This researcher chooses the effort to create data in the initial stage of the model, the number of papers written based on the data and whether and when to share the data with the research community. These decisions depend on the data availability policy of journals. Under No Policy, i.e., the status quo of most journals in economics, the creator of data can freely choose whether and when to share the data and whether to write one or two papers using the data. Following Dasgupta and David (1994), who emphasize that an institutional response is necessary to overcome the gap between researchers' demand for data sharing and its voluntary provision, we consider a second policy type, the First Paper Policy. The leading example is the data availability policy of the *American Economic Association,* which forces the creator of data to share the data after the first publication so that the data is available to the other researcher.

The creator of the data has a strong incentive to protect the competitive advantage associated with the data, as she might benefit from using the data for subsequent research and keep the data secret until their private value is fully exploited (Anderson et al., 2008; Haeussler, 2011; Haeussler et al., 2014; Stephan, 1996). If the data is kept private, another researcher cannot use it for replication or for subsequent research. The model endogenizes the decision to strategically delay the time of publication and thus the disclosure of data under the First Paper Policy. We derive exact conditions for positive welfare effects of mandatory data disclosure. However, we also find that the transition to mandatory data disclosure has negative welfare properties if it induces the data-creating researcher to delay her submissions strategically. In general, the welfare effects depend on the cost to generate data, the impatience of a researcher to publish (discount factor) and the additional value created by sharing data both for the data-creating researcher and the competing researcher.

Several top-tier economics journals have recently adopted mandatory data disclosure policies. However, the vast majority of economics journals either do not have a data availability policy or are reluctant to enforce it (McCullough, 2009; McCullough and Vinod, 2003). In addition, Andreoli-Versbach and Mueller-Langer (forthcoming) provide evidence for the status quo in economics with respect to voluntary data sharing using 488 observations randomly taken from researchers' academic webpages and public data repositories. Out of the sample, 435 researchers (89.14%) neither have a data and code section nor indicate whether and where their data is available. 8.81% of researchers share some of their data whereas only 2.05% fully share. Also, Anderson et al. (2008) suggest that authors generally hesitate to share their data and code despite their pre-publication commitment to provide this infor-

mation. This may suggest that editors, referees and readers are confident that the empirical results presented in the papers are always credible and robust. Unfortunately this is not the case (Lacetera and Zirulia, 2011). Empirical economics articles can often not be replicated (Dewald et al., 1986; McCullough et al., 2006 and 2008). This raises great concerns regarding the credibility and reliability of empirical work.

In a recent and heavily debated[2] article Angrist and Pischke (2010, p. 3) argue that better research design and the consequent causal interpretation of the regression coefficients "is taking the con out of econometrics". Even though the identification strategy is essential for good empirical work, without the possibility of replication and extension, which is provided by sharing the data and codes, doubts[3] about the credibility of empirical work remain.

In addition, researchers who do subsequent research have to develop the data themselves if the data are not shared, resulting in a (socially wasteful) duplication of research efforts. Open access to research data may on the one hand increase the credibility of authors as their work might be replicated by others and on the other hand facilitates new research as both data and codes would be readily available (McCullough and Vinod, 2003).[4] Even though most scholars agree that there are benefits of data sharing this view is not unanimous. A common objection to making data and code available is offered by Moffitt (2007) and McCullough (2009, p. 124):

> "Economists call the 'patent' problem the problem that those who put the effort into constructing a data set and writing programs (months of work) have the right to use it for further research for X years."

The remainder of the paper is organized as follows. In Section 2 we set up a simple two-player three-stage model of data disclosure. Section 3 analyzes the effects of different data disclosure policies on optimal effort and data sharing behaviour. In Section 4 we study the ambiguous welfare effects of mandatory data disclosure. Section 5 provides policy recommendations and Section 6 concludes our study.

[2]The replies to this article can be found here: http://econ-www.mit.edu/faculty/angrist/publications (last accessed June 6 2014).

[3]As McCullough states it in Glenn (2008): "If they [the authors] are not making their data and code available, then I have to think that they have something to hide. [...] There is a lot of nonreplicable research published in economics. We need to change the profession so that readers can expect that there actually does exist data and code that will reproduce the published results. Right now, at many journals, we cannot expect that."

[4]For example, in computational science, the data and the code are considered by peers as the real scientific contribution (Buckheit and Donoho, 1995). Claerbout's Principle states that: "An article about computational science in a scientific publication is not the scholarship itself, it is merely advertising of the scholarship. The actual scholarship is the complete software development environment and the complete set of instructions which generated the figures" (Buckheit and Donoho, 1995, p. 59).

2 A Simple Model of Data Disclosure

This model analyzes the optimal effort choices of a researcher, R, to generate novel data and to share the data produced with the scientific community, which is represented by a second researcher, C, who may use the data for subsequent research. We consider a three-stage model, $t = 0, 1, 2$, where the incentives to share depend on two factors. First, data disclosure may increase the value of a published article for R, e.g., it may increase the credibility of the article. Second, disclosure may change R's personal value of the data after the last publication. For instance, the loss of control of disclosed data may lead to flawed interpretation of the data, which may negatively affect R's reputation and thus her personal value of the data after the last publication. We study two data availability policies of journals. Under No Policy (henceforth NP), R can freely choose whether and when to share the data and whether to write one or two papers using the data. In contrast, the First Paper Policy (FPP) forces R to share the data after the first publication so that the data is available to C. For simplicity, we consider a maximum number of two journal publications that a single researcher can achieve by using the same data set. We assume that the marginal benefit from the re-use of the data is decreasing and that the returns to publication are diminishing with increased output (Tuckman and Leahey, 1973). Under both policies, R chooses effort e_0 to create a data set in the initial stage. The quality and the value of the data set and the publication increase in this initial effort. Thereby, it also affects R's decision to disclose the data and the quality of published articles which are derived from the data. R's effort to publish is given by e_1 in the first stage and by e_2 in the second stage. As Figure 1 illustrates, there are six possibilities, four under NP and two under FPP. Under NP, R may choose to (1) publish one paper and share in $t = 1$ (*1-Share*), (2) publish two subsequent papers in $t = 1$ and $t = 2$ and share in $t = 2$ (*2-Share*), (3) write one paper in $t = 1$ and never share (*1-Never*) and (4) publish two subsequent papers and never share (*2-Never*). Under FPP, R may (5) choose to publish the first paper in $t = 1$ while being forced to share the data. In stage 2, C will then choose effort x_2 and publish a paper using R's data (*No Strategic Delay*). (6) R may strategically delay the first publication in order to evade the forced disclosure of data and publish two papers in $t = 2$ (*Strategic Delay*). By "strategic delay" we mean that a researcher does not submit her papers one after another but all at the same time after the completion of the second paper in order to maintain the exclusive use of the data.

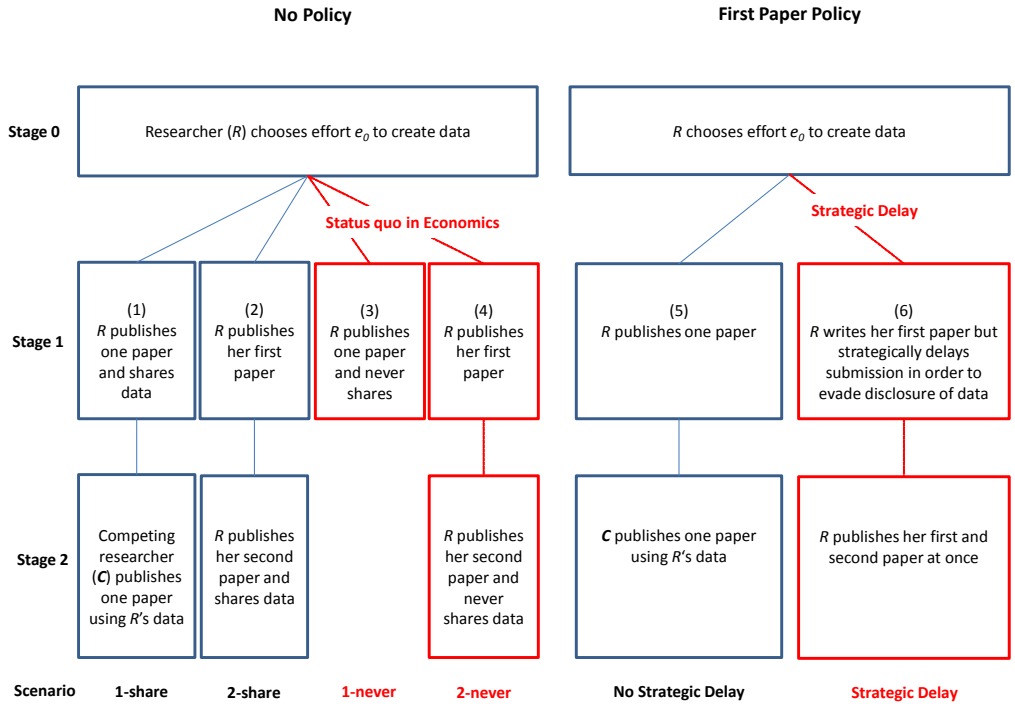

Figure 1: Stages of the model

Figure 1 illustrates a set of six reasonable publication strategies for R under the two data policies under study. For instance, R's decision to write one or two papers depends on the value of the data for subsequent use. In addition, R's decision to disclose or withhold data depends on her personal value of the data after the last publication. For instance, data disclosure may increase R's personal value of the data, e.g., due to subsequent use in data-journals. However, it may also be the case that the data has a higher value for R if it remains her private information. For instance, loss of control associated with data sharing can result in data falsification or flawed interpretation that may negatively affect R's reputation (Costello, 2009; Perrino et al., 2013). Thus, even though we assume that data disclosure always has a positive effect on the value of the published article that is based on this data, withholding the data after the final publication may be the optimal strategy for R in our model.

Let v be the value of a research idea that is generated by the scientific community based on R's data. Let $\alpha > 1$ be a creativity coefficient that measures the superior creativity of the scientific community as compared to the single researcher R. This is an important assumption as it implicitly states that the social optimum could be reached if the research was done by C. The intuition for this assumption is that C would achieve the highest value of research when he initially creates the data and writes both articles. In our model, however,

R initially produces the data and writes (at least) the first paper. For instance, we assume that R has the initial research idea together with exclusive access to data sources that give him a head start in the creation of data. Following Fudenberg et al. (1983)[5], we argue that this head start enables R to preempt C from generating data (as well as publication as long as the data are private information). A time coefficient, $c > 1$, indicates that the value of the second idea based on the data is lower than that of the first idea. Let $\frac{v}{\alpha}$ and $\frac{v}{c \cdot \alpha}$ be the value of the idea for R based on the data in stage 1 and 2, respectively. Henceforth, we normalize v to one. The higher c is, the stronger does the value of the idea based on the data devaluate. Let $\frac{1}{c}$ be the value of the idea based on the data for C in the second stage. As $\alpha > 1$, $\frac{1}{c} > \frac{1}{c \cdot \alpha}$. Thus, data availability creates positive spillovers as it enables C to publish an article of higher value as compared to R's second article under non-disclosure.

The cost functions are given by $c(e_t) = \frac{1}{2}e_t^2$ for R and $c(x_2) = \frac{1}{2}x_2^2$ for C, i.e., we assume increasing marginal costs of both data and article creation. Without data no research occurs and players get $U_{i,k_j^m} = 0$ with players $i = R, C$, policies $k = NP, FPP$, number of papers published $j = 1, 2$ and (in case that $k = NP$) R's sharing decision $m = share, never$. Welfare is given by $W_{k_j^m} = U_{R,k_j^m} + U_{C,k_j^m}$.

We will derive the optimal effort levels both for the data and paper creation. Then we will compare the total utility generated by R and C under different data disclosure policies in order to evaluate the policy implications of each regulation. We start by describing the setup of the model and the interaction between R and C under NP and under FPP. Then we will evaluate changes in the welfare function caused by a policy change from NP to FPP.

2.1 No Policy

In stage 0, R chooses effort e_0 to create the data. In stage 1, R chooses effort e_1 to publish the first paper that elaborates on the data and decides whether to share the data with the scientific community. In stage 2, there are two possibilities. First, if R does not share data in stage 1, she chooses effort e_2 to publish a second paper with the same data and whether to share the data with the scientific community. Second, if R shares the data in stage 1, C chooses effort x_2 to publish a paper by using R's data. R benefits from keeping the data private information. This exclusive use may increase R's ex ante incentives to exert effort on the data creation and preparation, which is an essential part of empirical economics. The higher R's initial effort to create the data, the higher are her benefits from the journal article that is based on the data. However, the scientific community, C, would benefit if R shares

[5] Fudenberg et al. (1983) analyze a patent race in which two firms compete for a preliminary discovery in the first stage and for a patent in the second stage. They show that the first entrant can preempt the follower's entering the race if the former has a sufficiently large head start.

the data. The benefits consist of the replicability of the results, further research without the costs of data duplication and better research ideas and methods as the data and codes are shared among a large community. The fundamental difference between the two researchers is that R has a data set which would be valuable for C, but C cannot influence R's decision when to share the data as there is no market for the exchange of data. The assumption that there is no market for data is very plausible in economics. The only market for data in economics is between institutions and researchers but not between researchers. Thus there is no monetary incentive to share the data and an efficient allocation of efforts not achievable (Coase, 1960).

Researcher publishes one paper and shares at $t = 1$ (1-Share): Appendix A (B) provides all calculations for the NP (FPP) scenario. R's optimization problem under *1-Share* is

$$\delta \left(e_0 e_1 \frac{1}{\alpha} (1 + \kappa) - \frac{1}{2} e_1^2 \right) + \delta^2 \gamma_D e_0 - \frac{1}{2} e_0^2 \longrightarrow \max, \tag{1}$$

where δ^t is the discount factor at time t which measures R's impatience to publish. For instance, this impatience arises from the researcher's ambition to push her academic career by obtaining tenure through publication (Dasgupta and David, 1994; Gans and Stern, 2010). To illustrate, for $\delta = 1$, R is indifferent between a publication today and a publication tomorrow. κ represents the change of the value of an article due to the disclosure of the data. We assume that κ is positive and not too large, $0 < \kappa \leq \frac{1}{2}$. For instance, suppose that the availability of the data used in an article is perceived by peers as a signal of quality (Dasgupta and David, 1994). In that sense, data sharing may be seen as a means to disclose R's private information on the quality of her empirical article. Thus, researchers of higher quality may be more likely to share their data voluntarily in order to signal the robustness of their results (Andreoli-Versbach and Mueller-Langer, forthcoming; Feigenbaum and Levy, 1993). In addition, empirical articles for which the applied data is publicly available may generate more citations. Let γ_D (γ_{ND}) measure the value of the data for R after the last publication if R discloses (does not disclose) the data and define $\xi = \frac{\gamma_D}{\gamma_{ND}}$. For instance, the data set might contain information that is valuable for non- or semi-academic projects, such as expert or policy advice, that do not necessarily lead to further scientific publications. Second, the data set itself might be cited in subsequent literature as it provides, for instance, a widely applicable index. In addition, a data set could become valuable if data journals are later established in economics. In these cases, the remaining personal value of the data after the last publication depends on whether the data are public or private information. The lower ξ the lower are *ceteris paribus* R's incentives to share data. The higher the scope for individual use of the data, the higher will be R's incentives to keep the data. The first

term in (1) is R's discounted net value from publishing in $t = 1$. It is given by the difference between the individual gross value from publishing and the effort cost to write the paper. The individual gross value from publishing depends positively on the effort to create the data, the effort to write the paper and κ. The second term is R's discounted value of the disclosed data set, which depends positively on the effort to create it. The third term describes the effort cost to create the data.

C's optimization problem under *1-Share* depends on R's optimal effort to create data, $e^*_{0,NP_1^{share}}$, and is given by:

$$\delta^2 \left(e^*_{0,NP_1^{share}} x_2 \frac{1}{c} - \frac{1}{2} x_2^2 \right) \longrightarrow \max . \tag{2}$$

Hence, C maximizes his discounted net value from publishing in $t = 2$, i.e. the gross value from publishing in $t = 2$ minus the effort cost to write the paper. We obtain the following optimal efforts: $e^*_{0,NP_1^{share}} = \frac{\delta^2 \gamma_D}{1-\delta\left(\frac{1}{\alpha}\right)^2 (1+\kappa)^2}$, $e^*_{1,NP_1^{share}} = e^*_{0,NP_1^{share}} \cdot \frac{1}{\alpha}(1+\kappa)$ and $x^*_{2,NP_1^{share}} = e^*_{0,NP_1^{share}} \cdot \frac{1}{c}$. Henceforth, we only consider positive optimal efforts and assume $1 > \delta \left(\frac{1}{\alpha}\right)^2 (1+\kappa)^2$.[6] Optimal efforts in all stages are increasing in γ_D and κ. The higher the value of the article and the data due to data disclosure, the higher is R's initial effort to create the data, $e^*_{0,NP_1^{share}}$. Optimal efforts chosen by C and R in stages 1 and 2 then depend on R's initial effort of data creation. The higher the quality of the data, the higher is the benefit from an article and thus the higher are the efforts to publish it. We obtain the following utilities for R and C, respectively: $U^*_{R,NP_1^{share}} = \frac{1}{2} e^*_{0,NP_1^{share}} \cdot \delta^2 \gamma_D$ and $U^*_{C,NP_1^{share}} = \frac{1}{2} (e^*_{0,NP_1^{share}})^2 \delta^2 \left(\frac{1}{c}\right)^2$. R's utility depends positively on the remaining personal value of the data after the last publication. It also increases if the positive κ-effect on the value of the article due to the disclosure of the data is more pronounced. C's utility increases if, at optimum, R increases his effort to create the data. It decreases if the devaluation of the value of the research idea based on R's data over time is more pronounced. Overall welfare is given by $W^*_{NP_1^{share}} = U^*_{R,NP_1^{share}} + U^*_{C,NP_1^{share}} = \frac{1}{2}\delta^2 (e^*_{0,NP_1^{share}} \cdot \gamma_D + (e^*_{0,NP_1^{share}})^2 \left(\frac{1}{c}\right)^2)$. It is rising in γ_D and κ and decreasing in c. Overall welfare is also decreasing in the inferiority of R's research idea as compared to C's research idea, which is measured by the creativity coefficient α. Note that, as the data is shared after the first stage, C publishes the second article. Sharing data after the first stage creates positive spillovers as data now can be used to extract the highest value of subsequent research.

[6] See also Proposition 0 below.

Researcher publishes two subsequent papers and shares at $t = 2$ (2-Share): R's optimization problem under *2-Share* is

$$\delta \left(e_0 e_1 \frac{1}{\alpha} (1 + \delta \cdot \kappa) - \frac{1}{2} e_1^2 \right) + \delta^2 \left(e_0 e_2 \frac{1}{c \cdot \alpha} (1 + \kappa) - \frac{1}{2} e_2^2 \right) + \delta^3 \gamma_D e_0 - \frac{1}{2} e_0^2 \longrightarrow \max . \quad (3)$$

The first term in (3) is the discounted net value of publishing in $t = 1$. Therein, the gross value of publishing is given by $e_0 e_1 \frac{1}{\alpha} (1 + \delta \cdot \kappa)$ and effort cost are given by $\frac{1}{2} e_1^2$. Note that in contrast to *1-Share*, the positive κ-effect is discounted under *2-Share* as data availability develops its positive effect on the value of the first article at a later stage. The second term in (3) is the discounted net value of publishing in $t = 2$. Therein, the gross value of publishing is given by $e_0 e_2 \frac{1}{c \cdot \alpha} (1 + \kappa)$. It accounts for the devaluation of the research idea over time. The third term is R's discounted value of the disclosed data set. The last term describes the effort cost to create the data. Again, we only consider positive optimal efforts, $1 > \delta \left(\frac{1}{\alpha} \right)^2 (1 + \delta \cdot \kappa)^2 + \delta^2 \left(\frac{1}{c \cdot \alpha} \right)^2 (1 + \kappa)^2$. We obtain the following optimal efforts: $e^*_{0,NP_2^{share}} = \frac{\delta^3 \gamma_D}{1 - \delta \left(\frac{1}{\alpha} \right)^2 (1+\delta \cdot \kappa)^2 - \delta^2 \left(\frac{1}{c \cdot \alpha} \right)^2 (1+\kappa)^2}$, $e^*_{1,NP_2^{share}} = e^*_{0,NP_2^{share}} \cdot \frac{1}{\alpha} (1 + \delta \cdot \kappa)$ and $e^*_{2,NP_2^{share}} = e^*_{0,NP_2^{share}} \cdot \frac{1}{c \cdot \alpha} (1 + \kappa)$. We obtain the following utility of R, which is equal to overall welfare as C does not generate any utility in this case: $U^*_{R,NP_2^{share}} = W^*_{NP_2^{share}} = \frac{1}{2} e^*_{0,NP_2^{share}} \cdot \delta^3 \gamma_D$. Overall welfare increases in γ_D and κ and decreases in c and α.

Researcher publishes one paper in $t = 1$ and never shares (1-Never): The optimization problem of R is given by

$$\delta \left(e_0 e_1 \frac{1}{\alpha} - \frac{1}{2} e_1^2 \right) + \delta^2 \gamma_{ND} e_0 - \frac{1}{2} e_0^2 \longrightarrow \max . \quad (4)$$

In contrast to *1-Share*, there is no positive κ-effect in the first term of (4). The second term is R's discounted value of the undisclosed data set. We obtain the following optimal efforts: $e^*_{0,NP_1^{never}} = \frac{\delta^2 \gamma_{ND}}{1 - \delta \left(\frac{1}{\alpha} \right)^2}$, $e^*_{1,NP_1^{never}} = e^*_{0,NP_1^{never}} \frac{1}{\alpha}$. We obtain the following utility (overall welfare): $U^*_{R,NP_1^{never}} = W^*_{NP_1^{never}} = \frac{1}{2} e^*_{0,NP_1^{never}} \cdot \delta^2 \gamma_{ND}$. Note that C does not generate any utility in this case. Overall welfare increases in γ_{ND} and decreases in α.

Researcher publishes two subsequent papers and never shares (2-Never): R's optimization problem under NP if she never shares[7] is

$$\delta \left(e_0 e_1 \frac{1}{\alpha} - \frac{1}{2} e_1^2 \right) + \delta^2 \left(e_0 e_2 \frac{1}{c \cdot \alpha} - \frac{1}{2} e_2^2 \right) + \delta^3 \gamma_{ND} e_0 - \frac{1}{2} e_0^2 \longrightarrow \max . \quad (5)$$

[7]Note that if R does not share the data immediately after her last publication, she will never share.

We obtain the following optimal efforts: $e^*_{0,NP_2^{never}} = \frac{\delta^3\gamma_{ND}}{1-\delta\left(\frac{1}{\alpha}\right)^2-\delta^2\left(\frac{1}{c\cdot\alpha}\right)^2}$, $e^*_{1,NP_2^{never}} = e^*_{0,NP_2^{never}} \cdot \frac{1}{\alpha}$ and $e^*_{2,NP_2^{never}} = e^*_{0,NP_2^{never}} \cdot \frac{1}{c\cdot\alpha}$. Note that $e^*_{0,NP_2^{never}} > 0$ when $e^*_{0,NP_2^{share}} > 0$. We obtain the following utility (overall welfare): $U^*_{R,NP_2^{never}} = W^*_{NP_2^{never}} = \frac{1}{2}e^*_{0,NP_2^{never}} \cdot \delta^3\gamma_{ND}$. Note that C does not generate any utility in this case. From the optimal efforts under the four NP scenarios, we obtain:

Proposition 0: Positive efforts condition *All optimal efforts are well-defined and positive for all* $\delta \in (0,1]$ *if* $(1+\kappa)^2 < \frac{c^2}{1+c^2}\alpha^2$.

All proofs are provided in Appendix C. However, straightforward proofs are omitted. We now analyze the conditions under which R does not share her data.

Proposition 1: Never Share *Under NP the Researcher can decide not to share only if* $\gamma_D < \gamma_{ND}$, i.e., if $\xi = \frac{\gamma_D}{\gamma_{ND}} < 1$. *(i) She will always choose not to share if* ξ *is sufficiently small, i.e.* $\xi \leq \xi_l$, *where* $\xi_l = \xi_l(\kappa)$ *depends negatively on* κ. *(ii) For* $\xi_l < \xi < 1$ *she will choose not to share if she is impatient enough, i.e.* $\delta \leq \delta_h$, *where* $\delta_h = \delta_h(\kappa, \xi)$ *depends nonpositively both on* κ *and* ξ.

(i) Intuitively, R does not have an incentive to share data if the remaining personal value of the undisclosed data after the last publication by far exceeds the personal value of the disclosed data. In this case, the disclosure-driven decrease in the remaining value of the data more than outweighs the positive "κ-effect" associated with the increase in the value of an article due to data disclosure. Note that (i) illustrates the status quo in economics as only very few economists voluntarily share their data with the scientific community (Andreoli-Versbach and Mueller-Langer, forthcoming). (ii) However, this result does not hold true in general if the above-stated decrease in the remaining value of the data is sufficiently low, i.e. $\xi > \xi_l$. Then, the level of patience comes into play and only a sufficiently impatient researcher will keep the data private information. In this case, the (strongly) discounted "κ-effect" does not outweigh the disclosure-driven decrease in the remaining value of the data.

Finally, note that, for $\xi \geq 1$, *1-Never* (*2-Never*) would be strictly dominated by *1-Share* (*2-Share*).

2.2 First Paper Policy

To simplify matters we omit labels for the obvious number of papers published, $j = 1, 2$, and R's sharing decision, $m = share, never$, in efforts, utilities and welfare under FPP.

First Paper Policy without Strategic Delay: Under FPP without strategic delay (NSD), R publishes the first paper in $t = 1$. She is required to share the data and does not compete with C in $t = 2$ (due to the competitive advantage that is captured in C's superior creativity). R's optimization problem is

$$\delta \left(e_0 e_1 \left(\frac{1}{\alpha} \right) (1 + \kappa) - \frac{1}{2} e_1^2 \right) + \delta^2 \gamma_D e_0 - \frac{1}{2} e_0^2 \longrightarrow \max . \tag{6}$$

R's maximization problem under NSD is the same as under *1-Share*. The main difference, however, is that she is forced to make the data available under FPP upon publication of the first paper. FPP is detrimental to R as it reduces her set of possible publication strategies.

The competitor C uses the data and publishes a paper in $t = 2$. C's optimization problem is

$$\delta^2 \left(e^*_{0,FPP_{NSD}} \cdot x_2 \frac{1}{c} - \frac{1}{2} x_2^2 \right) \longrightarrow \max . \tag{7}$$

The outcomes under NSD and *1-Share* are identical for $T = 0, 1, 2$ and $i = R, C$, that is, $e^*_{T,FPP_{NSD}} = e^*_{T,NP_1^{share}}$, $x^*_{2,FPP_{NSD}} = x^*_{2,NP_1^{share}}$, $U^*_{i,FPP_{NSD}} = U^*_{i,NP_1^{share}}$ and $W^*_{FPP_{NSD}} = W^*_{NP_1^{share}}$.

First Paper Policy with Strategic Delay: In contrast to the previous section, R may choose to delay her publications in order to fully exploit her data under FPP. In this case, she will incur the costs of completing the first and second paper in the first and second period, respectively. However, she will realize the total benefits associated with the publication of two papers in the second period, as she strategically delays publication in order to keep the data private. Under FPP with strategic delay (henceforth, SD), R does not publish in $t = 1$ but publishes the first and the second paper together in $t = 2$. R's optimization problem is

$$\delta \left(\delta e_0 e_1 \frac{1}{\alpha} (1 + \kappa) - \frac{1}{2} e_1^2 \right) + \delta^2 \left(e_0 e_2 \frac{1}{c \cdot \alpha} (1 + \kappa) - \frac{1}{2} e_2^2 \right) + \delta^3 \gamma_D e_0 - \frac{1}{2} e_0^2 \longrightarrow \max . \tag{8}$$

To understand R's optimization problem under SD, it is constructive to compare it with R's optimization problem under 2-share as given by (3). The second, third and fourth term is identical under both scenarios. The first term, however, is different if R is not perfectly patient, i.e., $\delta < 1$. In contrast to *2-Share*, R realizes the net value of the publication of the first paper at a later stage under SD. Under *2-Share*, the first paper is published right after completion in the first stage whereas its publication is postponed to the second stage under SD. We obtain the following optimal efforts: $e^*_{0,FPP_{SD}} = \frac{\delta^3 \gamma_D}{1 - \delta^3 \left(\frac{1}{\alpha} \right)^2 (1+\kappa)^2 - \delta^2 \left(\frac{1}{c \cdot \alpha} \right)^2 (1+\kappa)^2}$, $e^*_{1,FPP_{SD}} = \delta \cdot e^*_{0,FPP_{SD}} \cdot \frac{1}{\alpha} (1 + \kappa)$ and $e^*_{2,FPP_{SD}} = e^*_{0,FPP_{SD}} \cdot \frac{1}{c \cdot \alpha} (1 + \kappa)$. We obtain the following optimal utility (overall welfare): $U^*_{R,FPP_{SD}} = W^*_{FPP_{SD}} = \frac{1}{2} e^*_{0,FPP_{SD}} \delta^3 \gamma_D$. Note that C does

not generate any utility in this case. Overall welfare increases in γ_D and κ and decreases in c and α.

3 Analysis and Comparison of No Policy and First Paper Policy

We analyze the effect of the transition from NP to FPP on R's efforts to create data. We obtain:

Proposition 2: Change in efforts under FPP *Transition to FPP (i) reduces efforts to create data, (1) if ξ is sufficiently small, i.e. $\xi < \xi_{l_e} < 1$, where $\xi_{l_e} = \xi_{l_e}(\kappa)$ depends negatively on κ, or (2) if $\xi_{l_e} < \xi < 1$ and the Researcher is impatient enough, i.e. $\delta < \delta_{h_e}$, where $\delta_{h_e} = \delta_{h_e}(\kappa, \xi)$ depends negatively on both κ and ξ, or (3) if $\xi > \xi_h$ and the Researcher is very patient, i.e. $\delta > \delta_{h_{e_1}} := \max\{\delta_{h_1}, \delta_{h_{SD}}\}$, where $\delta_{h_{e_1}} = \delta_{h_{e_1}}(\kappa)$ depends negatively on κ. (ii) does not influence efforts to create data if the Researcher chooses to share after one publication, which is the case when $\xi \geq 1$ and the Researcher is not very patient, i.e. $\delta \leq \delta_{h_1}$, where $\delta_{h_1} = \delta_{h_1}(\kappa)$ depends negatively on κ, or when $\xi_h \leq \xi \leq 1$ and $\delta_h \leq \delta \leq \delta_{h_1}$. (iii) increases efforts to create data (1) if $\xi_{l_e} < \xi < 1$ and $\delta_{h_e} < \delta < \min\{\delta_h, \delta_{h_{SD}}\}$, or (2) if $\xi_{ul_e}(\delta) < \xi < \xi_{h_1}$ and $\delta > \max\{\delta_{0_1}, \delta_{h_{SD}}, \delta_{h_1}\}$.*

The economic intuition behind these results is the following. (i)(1) Intuitively, if ξ is sufficiently low, i.e. sharing the data would harm R significantly, the transition to FPP reduces her incentives to create data in the first place independently of R's publication choice under NP. (i)(2) If *1-Never* is optimal for R under NP, i.e. ξ is sufficiently low, the transition to FPP reduces R's incentives to create the data. Then, the obligation to share reduces her returns from the data. Note, however, that the transition does not induce R to delay submission strategically in this case. (i)(3) Whenever R chooses to delay submission strategically under FPP and to share after two publications under NP, the transition to FPP reduces R's efforts to create data. Intuitively, both scenarios (*2-Share* under NP and SD under FPP) lead to the same outcome in the sense that the data are available after two publications. However, R generates a lower (discounted) utility from the first publication under FPP with SD as compared to *2-Share* under NP.

Under the conditions specified in (ii), *1-Share* is optimal for R under NP and NSD is optimal under FPP. The transition to FPP does not influence R's effort to create data as she is indifferent between *1-Share* under NP and NSD under FPP.

(iii) Our analysis suggests that the transition to FPP potentially increases R's efforts to create data under the following conditions: R is sufficiently patient to publish and ξ is neither too low, i.e. data sharing is not too harmful, nor too high, i.e. R would not share data voluntarily under NP. However, the result that the transition to FPP may increase R's incentives to create data may seem counterintuitive as one may argue that the set of choices under FPP is a subset of the set of choices under NP. Note, however, that the transition to FPP does not make R better off in terms of utility. It forces R to adjust her optimal effort choice, which *ceteris paribus* turns out to be effort-increasing under the conditions specified in (iii).

We provide the exact conditions under which R finds it optimal to delay submission strategically under FPP in the following proposition.

Proposition 3: Strategic Delay *The Researcher will delay strategically under FPP if and only if she is sufficiently patient, i.e. $\delta \geq \delta_{hSD}$, where $\delta_{hSD} = \delta_{hSD}(\kappa)$ depends negatively on κ.*

Intuitively, if R is sufficiently patient, the additional benefits from the second publication more than outweigh the delayed realization of the benefits from the first paper under SD as compared to NSD. Regarding the interpretation of this proposition, note that junior researchers are typically more impatient than senior researchers because they are in need of a good publication soon in order to obtain a tenured position (Dasgupta and David, 1994; Gans and Stern, 2010). Hence, one may interpret Proposition 3 in the sense that junior researchers are *ceteris paribus* less likely to strategically delay their submission than senior (tenured) researchers. However, it is also reasonable to assume that the value of an additional publication in terms of career concerns is higher for (untenured) junior researchers than for (tenured) senior researchers for the same reason (Coupé et al., 2006; Siegfried and White, 1973; Tuckman and Leahey, 1973).[8] This would make SD relatively more attractive to young scholars. Hence, the question of whether junior researchers have lower incentives than senior researchers to delay strategically depends on whether the "impatience effect" dominates the countervailing "marginal publication effect" under the assumption that the rate of data depreciation does not differ between junior and senior researchers.

[8]For instance, writing about the career incentives of tenured professors, Stanford's tenured Professor Robert Hall (2009, p. 1) states: "Now that you have tenure, the number of papers you produce is amazingly irrelevant."

4 Welfare Analysis

We analyze the welfare properties of the transition to FPP. The precise conditions under which this transition has neutral, negative or positive welfare properties are given in the following proposition.

Proposition 4: Welfare effects *Transition to FPP (i) does not influence welfare if the Researcher chooses to share after one publication, which is the case when $\xi \geq 1$ and the Researcher is not very patient, i.e. $\delta \leq \delta_{h_1}$, where $\delta_{h_1} = \delta_{h_1}(\kappa)$ depends negatively on κ, or when $\xi_h \leq \xi \leq 1$ and $\delta_h \leq \delta \leq \delta_{h_1}$. (ii) (1) It increases welfare if $\xi_{l_w} < \xi \leq 1$ and $\delta_l < \delta < \delta_{h_0} := \min\{\delta_h, \delta_{h_{SD}}\}$, where $\xi_{l_w} = \xi_{l_w}(\kappa)$ depends negatively on κ and both $\delta_l = \delta_l(\kappa, \xi)$ and $\delta_{h_0} = \delta_{h_0}(\kappa, \xi)$ depend negatively both on κ and ξ, or (2) if $\xi_{l_3} < \xi < \xi_{h_1}$ and $\delta_{0_1} < \delta < \min\{\delta_{h_{SD}}, \delta_{0_0}\}$, where $\xi_{l_3}(\kappa)$ and $\xi_{h_1}(\kappa)$ depend negatively on κ. (iii) It reduces welfare when ξ is sufficiently small, i.e. $\xi < \xi_{l_1}$, where $\xi_{l_1} = \xi_{l_1}(\kappa)$ depends negatively on κ, or if the Researcher is very patient, i.e. $\delta > \delta_{h_{SD}}$.*

The economic intuition behind these results is the following. (i) *1-Share* under NP and NSD under FPP lead to the same outcome from a social welfare perspective. Both researchers are indifferent between NP and FPP.

(ii) (1) Under these conditions, R switches from *1-Never* under NP to NSD under FPP. The welfare gain due to the additional paper published by C under FPP more than outweighs the potentially negative impact of forced data sharing on R. (2) R switches from *2-Never* to NSD under FPP while ξ is sufficiently high.

(iii) These results specify the conditions under which the transition to FPP has negative welfare properties. First, the transition to FPP reduces welfare if it induces R to delay submission strategically. If R delays strategically, two negative effects emerge: There will be a delay in the time the scientific knowledge is made public and C cannot use the data to publish his (higher value) paper. Analytically, it can be shown that R's utility under the conditions which incentivize her to choose *2-Share* under NP always exceeds her utility under SD under FPP. In addition, for all combinations of ξ and δ for which *2-Share* is not chosen under NP, it is dominated by one of the other strategies under NP. Therefore, SD is not only dominated by *2-Share* but also by all other possible strategies under NP. Thus, the transition to SD always decreases R's utility and thus overall welfare independent of the researcher's strategy under NP. Second, for sufficiently low ξ, the transition to FPP reduces welfare if it forces R to switch from never sharing after the first or second publication under NP to not delaying strategically. In this case, the negative effect of the transition

to FPP on the constrained R more than outweighs its positive effect on C. Forced data sharing associated with the transition to FPP is not a recommendable policy option under the conditions specified in (iii). It is a necessary (but not sufficient) condition for positive welfare properties of forced data sharing that the implementation of FPP is complemented by other policies which deter strategic delay, such as career incentives.

Graphical Approximation of the Main Results Figure 2 illustrates R's choices and the effects of the transition from NP to FPP depending on the impatience rate, δ (vertical axis), and the ratio of data values after all publications in case of disclosure/nondisclosure, $\xi = \frac{\gamma_D}{\gamma_{ND}}$ (horizontal axis). Other model parameters are kept constant ($c = \alpha = 1.5$ and $\kappa = 0.2$).

The black lines divide the whole area into 4 zones, indicating R's choices under NP: R chooses *1-Never* in the bottom left zone, *2-Never* in the top left zone (see Proposition 1), *2-Share* in the top right zone and *1-Share* in the bottom right zone. The horizontal blue line denotes the border for R's choice under FPP: she chooses to delay strategically (SD) above the blue line and not to delay (NSD) below it. It represents $\delta_{h_{SD}}$ from P(roposition) 3.

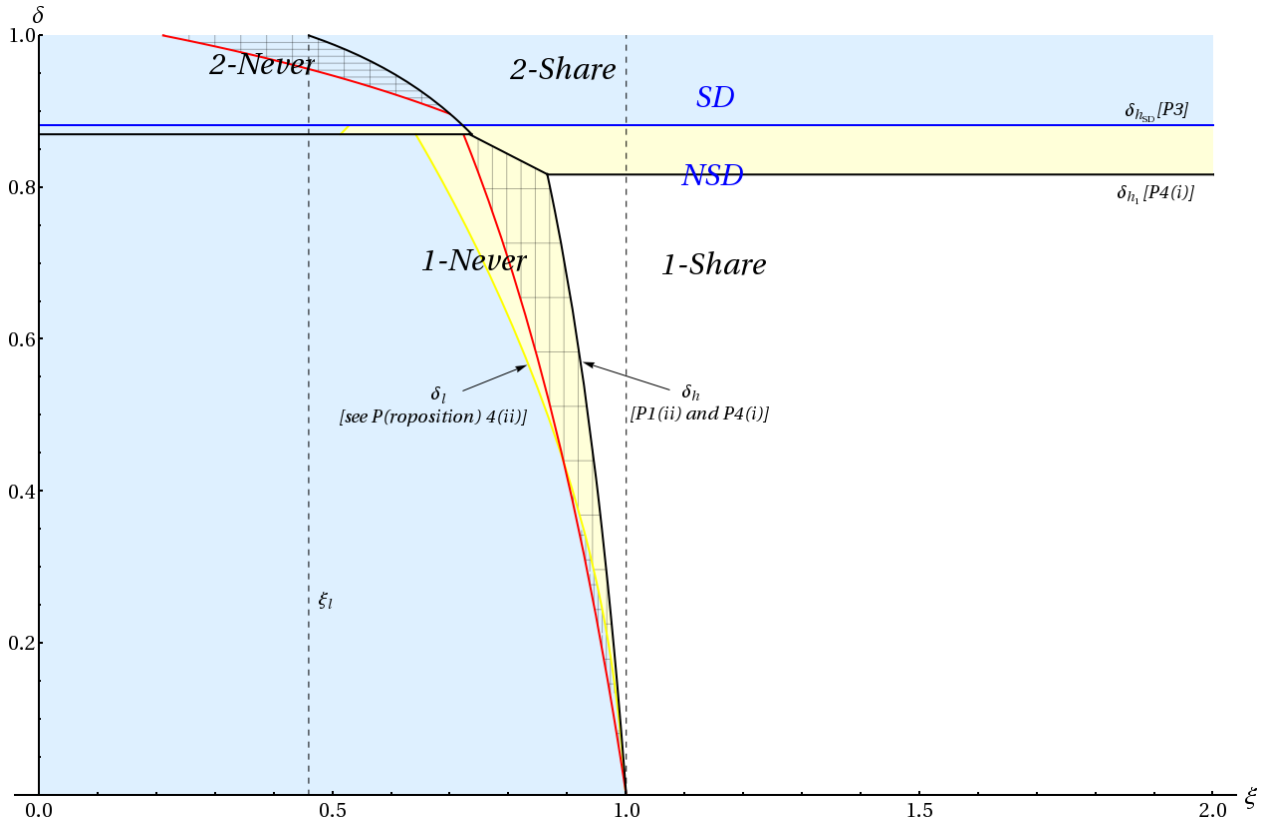

Figure 2: Researcher's choices and the effects of the transition from NP to FPP

The light blue area indicates those combinations of δ and ξ for which the transition from NP to FPP is welfare reducing. The light yellow area indicates combinations for which the transition to FPP increases welfare.[9] As for the white zone, which is exactly the zone where R chooses *1-Share* under NP, the transition to FPP does not change welfare (Proposition 4(i)). Finally, the checked areas (always bordered by red lines on one side) indicate where the transition to FPP increases efforts to create the data. Most notably, Figure 2 suggests that there is scope for a socially beneficial transition to FPP that increases the effort to create data, as indicated by the overlapping checked and light yellow areas. Efforts do not change after transition to FPP in the white zone and decrease everywhere else (Proposition 2). Strategic delay under FPP is unambiguously welfare reducing.

5 Policy Recommendations

The public availability of research data is an important issue from a research policy perspective (European Commission, 2012; OECD, 2007; US House of Representatives, 2007). Several journals and research-funding organizations have recently introduced data availability policies (Andreoli-Versbach and Mueller-Langer, forthcoming; ESRC, 2010; McCullough, 2009; NIH, 2003; NSF, 2011; Wellcome Trust, 2007). However, compiling and documenting data and code for subsequent use is costly and the data creator's benefits of shared data in the form of recognition or citations to the data are relatively low (Anderson et al., 2008; Costello, 2009).

On the basis of our findings and the fact that an institutional response is highly needed to increase data sharing (Dasgupta and David, 1994) we believe that mandatory data disclosure policies are necessary. However, they may be implemented along with higher incentives to share data in order to prevent the creator of the data from choosing to delay her submissions and decrease the effort in data generation. If such incentives are not implemented younger researchers who have a higher impatience to publish and possibly have larger cost of data creation because of inexperience and lack of research assistants might have to carry the greatest burden from mandatory data disclosure.

5.1 Data Journals and Citations

As the social value of knowledge increases by sharing the data one variable which can be changed by policy makers is the value of the data for their creator after the last publication if she discloses the data, γ_D (see Proposition 1). In particular, γ_D could be the increasing value

[9] See also the proof of Proposition 4 (ii) in Appendix C.

of data citations[10] on academic reputation.[11] In this respect, policy-makers and research funding agencies might incentivize and support the establishment of data journals. The establishment of data journals in economics may create a market for the exchange of data and increase the value of a data set for its creator and the availability of data.[12] This may also increase the quality of available data if data publications are subject to a peer-review process. Lessons could be learned from other disciplines. For instance, Nature (2013) has announced the launch of the online data journal *Scientific Data.* Moreover, the online-only *Geoscience Data Journal* has recently been launched (Allan, 2012). In addition, incentive schemes rewarding the production and documentation of data (Fienberg et al., 1985), such as a new standard of data citation (Altman and King, 2007), should be implemented so that sharing becomes more valuable to the original creator of data. Examples of recently established tools for data retention and citation are, among others, Dryad, Figshare and Zenodo. In general, if the individual academic value of data increases, i.e. due to data journals and increased and more valuable data citations, some researchers may have an incentive to specialize in the creation of data if they expect to have a competitive advantage, i.e. they have the necessary (financial) resources, knowledge and experience in the creation of data. This may have a positive effect on the overall quality of available data.

5.2 Journals for Replication Studies

As the publication market for replication studies in economics is limited, investing time and effort in writing a replication study is not an efficient use of a researcher's resources (Hamermesh, 1997; McCullough et al., 2006; Mirowski and Sklivas, 1991). However, the establishment of journals for replication studies may induce more researchers to engage in this field. The establishment of such journals could strengthen our so-called κ-effect of data sharing on the value of published articles. It would increase the ex ante incentives to produce correct results as authors will anticipate that their papers might be subject to the scrutiny of replicators. This would be beneficial for the academic community as incorrect results could be removed from the cumulative body of knowledge. The data creating author may also benefit ex post as replication may increase the credibility of her work. In addition, we hypothesize that replication studies could be particularly beneficial for junior researchers, i.e. PhD candidates or Master students on a research track, due to learning effects. For

[10] For example, information about the newly implemented data citation standard is available at http://thedata.org/citation.

[11] An example of where to share and find data is Dataverse Network, which offers a solution for publishing research data or accessing data from other researchers for replication studies. See http://thedata.org/book/use-and-share-data.

[12] See Gans and Stern (2010) for a similar discussion on the missing market for ideas in research.

instance, supervising professors may require their Master students or PhD candidates to actively contribute to such journals under their guidance. Replication studies would teach junior researchers the responsible treatment of data and put additional pressure on senior researchers to generate correct data in the first place (Feigenbaum and Levy, 1993; McCullough et al., 2006).

5.3 Possible Actions for Universities

First, universities could reward data sharing behavior through additional research grants or by positively considering sharing in a researchers' career. For instance, status- and priority-based rewards such as tenure and prizes for outstanding achievements may incentivize researchers to disclose their data (Dasgupta and David, 1994; Fienberg et al., 1985; Gans and Stern, 2010; Mukherjee and Stern, 2009). Second, universities could promote data sharing through institutional assistance. According to Kim and Stanton (2012), the compilation, preparation and sharing of research data are perceived as costly, thereby preventing researchers from sharing their data. Reducing these obstacles could result in a higher net benefit of sharing data and help to generate a pro-data-sharing attitude. For instance, some universities have recently taken the lead in this respect by developing data repositories and structured guidance for the creation of data management plans (Andreoli-Versbach and Mueller-Langer, forthcoming).

6 Conclusion

We set up a simple model describing the incentives of a researcher to generate novel data, publish articles and share her data with the research community so that other researchers can use the data for subsequent empirical research. We compare two different policies. First, under "No Policy", the creator of data has complete freedom as when to share her data voluntarily. Second, under the "First Paper Policy", she is required to share the data immediately after her first publication.

The implementation of First Paper Policy may distort her incentives in two fundamental ways. First, she might strategically delay her submissions in order to continue publishing with the same data without making it available to others. Second, she might reduce her effort to create data as she would not be the full residual claimant of her data after the first publication. Strategic delaying would prevent other researchers from working on related subsequent research with the same data. The decision to delay the time of submission depends on three factors. The first of these is the impatience to publish of the data-creating

researcher (as given by the discount factor); the second is the value of the data, i.e. the more valuable the data the higher is the probability that the researcher will choose to delay her submissions. The final factor is how fast the value of the data depreciates as the number of publications increases.

Our welfare analysis suggests that a mandatory data disclosure policy has negative welfare properties if it induces the researcher to strategically delay the submission of subsequent papers. Under this scenario the research community would not be in a position to make use of the data when it is still valuable for further research. Mandatory data disclosure may only be welfare enhancing if researchers have no incentives to postpone the time of their publication and if the positive effect of data availability outweighs the negative effect associated with reduced efforts to create data. We conclude that the implementation of mandatory data sharing rules should be complemented by other policies that deter strategic delay, such as career incentives, and increase the stand-alone value of academic data, such as new standards for data citation and the establishment of data journals. We also recommend the establishment of journals for replication studies in economics.

As a first idea for further research, we suggest relaxing the implicit assumption that the two researchers never work on independent papers (with the same topic) at the same time. This is particularly likely under the First Paper Policy and may result in a socially wasteful duplication of research efforts.

Finally, as a publication in a top-tier economics journal is highly desired by most researchers and might lead to promotion within a university, the benefits of a publication are likely to outweigh the cost of data sharing. It is questionable whether this result holds for medium-ranked journals with a data availability policy, where authors may choose competing journals with similar ranking but without a data availability policy.

7 Appendix

7.1 Appendix A: Optimal efforts, utility and welfare under NP

7.1.1 Researcher publishes one paper and shares at $t = 1$ (*1-Share*)

Optimal Efforts and Utility of Researcher R Under *1-Share*, R maximizes

$$U_{R,NP_1^{share}} = \delta \left(e_0 e_1 \frac{1}{\alpha} (1 + \kappa) - \frac{1}{2} e_1^2 \right) + \delta^2 \gamma_D e_0 - \frac{1}{2} e_0^2.$$

The first-order conditions (FOCs) are given by

$$\begin{aligned}\frac{\partial U_{R,NP_1^{share}}}{\partial e_0} &= 0 \Rightarrow \delta e_1 \frac{1}{\alpha}(1+\kappa) + \delta^2 \gamma_D = e_0(e_1), \\ \frac{\partial U_{R,NP_1^{share}}}{\partial e_1} &= 0 \Rightarrow e_0 \frac{1}{\alpha}(1+\kappa) = e_1(e_0).\end{aligned}$$

By plugging $e_1(e_0)$ in $e_0(e_1)$, we obtain $e^*_{0,NP_1^{share}} = \frac{\delta^2\gamma_D}{1-\delta\left(\frac{1}{\alpha}\right)^2(1+\kappa)^2}$. Then, by plugging $e^*_{0,NP_1^{share}}$ in $e_1(e_0)$, we obtain $e^*_{1,NP_1^{share}} = \frac{\delta^2\gamma_D\frac{1}{\alpha}(1+\kappa)}{1-\delta\left(\frac{1}{\alpha}\right)^2(1+\kappa)^2}$. By setting $e_1(e_0)$ in $U_{R,NP_1^{share}}$, we obtain:

$$\delta\frac{1}{2}e_0^2\left(\frac{1}{\alpha}\right)^2(1+\kappa)^2 + \delta^2\gamma_D e_0 - \frac{1}{2}e_0^2 = \frac{1}{2}\left(e_0^2\left(\delta\left(\frac{1}{\alpha}\right)^2(1+\kappa)^2 - 1\right) + 2\delta^2\gamma_D e_0\right).$$

Now plug in $e^*_{0,NP_1^{share}}$:

$$\begin{aligned}&\frac{1}{2}\left(\left(\frac{\delta^2\gamma_D}{1-\delta\left(\frac{1}{\alpha}\right)^2(1+\kappa)^2}\right)^2\left(\delta\left(\frac{1}{\alpha}\right)^2(1+\kappa)^2 - 1\right) + 2\delta^2\gamma_D\frac{\delta^2\gamma_D}{1-\delta\left(\frac{1}{\alpha}\right)^2(1+\kappa)^2}\right) \\ = \;& \frac{1}{2}\left(-\frac{\delta^4\gamma_D^2}{1-\delta\left(\frac{1}{\alpha}\right)^2(1+\kappa)^2} + \frac{2\delta^4\gamma_D^2}{1-\delta\left(\frac{1}{\alpha}\right)^2(1+\kappa)^2}\right) \\ \Leftrightarrow \;& U^*_{R,NP_1^{share}} = \frac{1}{2}\frac{\delta^4\gamma_D^2}{1-\delta\left(\frac{1}{\alpha}\right)^2(1+\kappa)^2} = \frac{1}{2}e^*_{0,NP_1^{share}}\cdot\delta^2\gamma_D.\end{aligned}$$

***C*'s Optimal Efforts and Utility** Under *1-Share*, C maximizes

$$U_{C,NP_1^{share}} = \delta^2\left(e^*_{0,NP_1^{share}}x_2\frac{1}{c} - \frac{1}{2}x_2^2\right).$$

The FOC leads to:

$$\frac{\partial U_{C,NP_1^{share}}}{\partial x_2} = 0 \Rightarrow e^*_{0,NP_1^{share}}\cdot\frac{1}{c} = x_2\left(e^*_{0,NP_1^{share}}\right).$$

By plugging $x_2\left(e^*_{0,NP_1^{share}}\right)$ in $U_{C,NP_1^{share}}$, we obtain:

$$\begin{aligned}
& \delta^2\left(e^*_{0,NP_1^{share}}\left(e^*_{0,NP_1^{share}}\cdot\frac{1}{c}\right)\frac{1}{c}-\frac{1}{2}\left(e^*_{0,NP_1^{share}}\cdot\frac{1}{c}\right)^2\right)\\
\iff\ & \delta^2\left(\left(e^*_{0,NP_1^{share}}\right)^2\left(\frac{1}{c}\right)^2-\frac{1}{2}\left(e^*_{0,NP_1^{share}}\right)^2\left(\frac{1}{c}\right)^2\right)\\
\iff\ & U^*_{C,NP_1^{share}}=\frac{1}{2}\left(e^*_{0,NP_1^{share}}\right)^2\delta^2\left(\frac{1}{c}\right)^2.
\end{aligned}$$

Welfare under NP if R writes one paper and shares at $t=1$ is given by:

$$\begin{aligned}
W^*_{NP_1^{share}} &= U^*_{R,NP_1^{share}}+U^*_{C,NP_1^{share}}\\
&\iff W^*_{NP_1^{share}}=\frac{1}{2}e^*_{0,NP_1^{share}}\cdot\delta^2\gamma_D+\frac{1}{2}\left(e^*_{0,NP_1^{share}}\right)^2\delta^2\left(\frac{1}{c}\right)^2\\
&\iff W^*_{NP_1^{share}}=\frac{1}{2}\delta^2\left(e^*_{0,NP_1^{share}}\cdot\gamma_D+\left(e^*_{0,NP_1^{share}}\right)^2\left(\frac{1}{c}\right)^2\right).
\end{aligned}$$

7.1.2 Researcher publishes two subsequent papers and shares at $t=2$ (*2-Share*)

Under *2-Share*, R maximizes

$$U_{R,NP_2^{share}}=\delta\left(e_0e_1\frac{1}{\alpha}\left(1+\delta\cdot\kappa\right)-\frac{1}{2}e_1^2\right)+\delta^2\left(e_0e_2\frac{1}{c\cdot\alpha}\left(1+\kappa\right)-\frac{1}{2}e_2^2\right)+\delta^3\gamma_De_0-\frac{1}{2}e_0^2.$$

The FOCs lead to:

$$\begin{aligned}
\frac{\partial U_{R,NP_2^{share}}}{\partial e_0} &= 0\Rightarrow\delta e_1\frac{1}{\alpha}\left(1+\delta\kappa\right)+\delta^2e_2\frac{1}{c\cdot\alpha}\left(1+\kappa\right)+\delta^3\gamma_D=e_0\left(e_1,e_2\right),\\
\frac{\partial U_{R,NP_2^{share}}}{\partial e_1} &= 0\Rightarrow e_0\frac{1}{\alpha}\left(1+\delta\kappa\right)=e_1\left(e_0\right),\\
\frac{\partial U_{R,NP_2^{share}}}{\partial e_2} &= 0\Rightarrow e_0\frac{1}{c\cdot\alpha}\left(1+\kappa\right)=e_2\left(e_0\right).
\end{aligned}$$

By plugging $e_1(e_0)$ and $e_2(e_0)$ in $e_0(e_1, e_2)$, we obtain

$$\begin{aligned} e_0 &= e_0\delta\left(\frac{1}{\alpha}\right)^2(1+\delta\kappa)^2 + e_0\delta^2\left(\frac{1}{c\cdot\alpha}\right)^2(1+\kappa)^2 + \delta^3\gamma_D \\ &\iff e_0 - e_0\delta\left(\frac{1}{\alpha}\right)^2(1+\delta\kappa)^2 - e_0\delta^2\left(\frac{1}{c\cdot\alpha}\right)^2(1+\kappa)^2 = \delta^3\gamma_D \\ &\iff e^*_{0,NP_2^{share}} = \frac{\delta^3\gamma_D}{1-\delta\left(\frac{1}{\alpha}\right)^2(1+\delta\kappa)^2 - \delta^2\left(\frac{1}{c\cdot\alpha}\right)^2(1+\kappa)^2}. \end{aligned}$$

By setting $e^*_{0,NP_2^{share}}$ in $e_1(e_0)$ and $e_2(e_0)$ we obtain:

$$\begin{aligned} e^*_{1,NP_2^{share}} &= \frac{\delta^3\gamma_D\frac{1}{\alpha}(1+\delta\kappa)}{1-\delta\left(\frac{1}{\alpha}\right)^2(1+\delta\kappa)^2 - \delta^2\left(\frac{1}{c\cdot\alpha}\right)^2(1+\kappa)^2}, \\ e^*_{2,NP_2^{share}} &= \frac{\delta^3\gamma_D\frac{1}{c\cdot\alpha}(1+\kappa)}{1-\delta\left(\frac{1}{\alpha}\right)^2(1+\delta\kappa)^2 - \delta^2\left(\frac{1}{c\cdot\alpha}\right)^2(1+\kappa)^2}. \end{aligned}$$

By setting $e_1(e_0)$ and $e_2(e_0)$ in $U_{R,NP_2^{share}}$, we obtain:

$$\begin{aligned} &\delta\frac{1}{2}e_0^2\left(\frac{1}{\alpha}\right)^2(1+\delta\kappa)^2 + \delta^2\frac{1}{2}e_0^2\left(\frac{1}{c\cdot\alpha}\right)^2(1+\kappa)^2 + \delta^3\gamma_D e_0 - \frac{1}{2}e_0^2 \\ = &\frac{1}{2}\left(\delta e_0^2\left(\frac{1}{\alpha}\right)^2(1+\delta\kappa)^2 + \delta^2 e_0^2\left(\frac{1}{c\cdot\alpha}\right)^2(1+\kappa)^2 + 2\delta^3\gamma_D e_0 - e_0^2\right) \\ = &\frac{1}{2}\left(e_0^2\left(\delta\left(\frac{1}{\alpha}\right)^2(1+\delta\kappa)^2 + \delta^2\left(\frac{1}{c\cdot\alpha}\right)^2(1+\kappa)^2 - 1\right) + 2\delta^3\gamma_D e_0\right). \end{aligned}$$

Now plug in $e^*_{0,NP_2^{share}}$:

$$\begin{aligned} &\frac{1}{2}\left(\begin{array}{c}\left(\frac{\delta^3\gamma_D}{1-\delta\left(\frac{1}{\alpha}\right)^2(1+\delta\kappa)^2-\delta^2\left(\frac{1}{c\cdot\alpha}\right)^2(1+\kappa)^2}\right)^2\left(-\left(1-\delta\left(\frac{1}{\alpha}\right)^2(1+\delta\kappa)^2-\delta^2\left(\frac{1}{c\cdot\alpha}\right)^2(1+\kappa)^2\right)\right) \\ +\left(\frac{2\delta^6\gamma_D^2}{1-\delta\left(\frac{1}{\alpha}\right)^2(1+\delta\kappa)^2-\delta^2\left(\frac{1}{c\cdot\alpha}\right)^2(1+\kappa)^2}\right)\end{array}\right) \\ \iff\ & U^*_{R,NP_2^{share}} = \frac{1}{2}\frac{\delta^6\gamma_D^2}{1-\delta\left(\frac{1}{\alpha}\right)^2(1+\delta\kappa)^2 - \delta^2\left(\frac{1}{c\cdot\alpha}\right)^2(1+\kappa)^2} = \frac{1}{2}e^*_{0,NP_2^{share}}\cdot\delta^3\gamma_D. \end{aligned}$$

7.1.3 Researcher publishes one paper in $t = 1$ and never shares (*1-Never*)

Under *1-Never*, R maximizes

$$U_{R,NP_1^{never}} = \delta\left(e_0e_1\frac{1}{\alpha} - \frac{1}{2}e_1^2\right) + \delta^2\gamma_{ND}e_0 - \frac{1}{2}e_0^2.$$

We obtain:

$$\begin{aligned}\frac{\partial U_{R,NP_1^{never}}}{\partial e_0} &= 0 \Rightarrow \delta e_1 \frac{1}{\alpha} + \delta^2 \gamma_{ND} = e_0\left(e_1\right), \\ \frac{\partial U_{R,NP_1^{never}}}{\partial e_1} &= 0 \Rightarrow e_0 \frac{1}{\alpha} = e_1\left(e_0\right).\end{aligned}$$

Plug $e_1\left(e_0\right)$ in $e_0\left(e_1\right)$:

$$e_0 = \delta\left(e_0 \frac{1}{\alpha}\right)\frac{1}{\alpha} + \delta^2\gamma_{ND} \Longleftrightarrow e^*_{0,NP_1^{never}} = \frac{\delta^2\gamma_{ND}}{1-\delta\left(\frac{1}{\alpha}\right)^2}.$$

Plug $e^*_{0,NP_1^{never}}$ in $e_1\left(e_0\right)$ to obtain $e^*_{1,NP_1^{never}}$:

$$e^*_{1,NP_1^{never}} = e^*_{0,NP_1^{never}}\frac{1}{\alpha}.$$

By plugging $e^*_{0,NP_1^{never}}$ and $e^*_{1,NP_1^{never}}$ in $U_{R,NP_1^{never}}$, we obtain the following utility and welfare:

$$\begin{aligned}&\delta\left(\frac{\delta^2\gamma_{ND}}{1-\delta\left(\frac{1}{\alpha}\right)^2}\frac{\delta^2\gamma_{ND}}{1-\delta\left(\frac{1}{\alpha}\right)^2}\left(\frac{1}{\alpha}\right)^2 - \frac{1}{2}\left(\frac{\delta^2\gamma_{ND}}{1-\delta\left(\frac{1}{\alpha}\right)^2}\frac{1}{\alpha}\right)^2\right) \\ &+\delta^2\gamma_{ND}\left(\frac{\delta^2\gamma_{ND}}{1-\delta\left(\frac{1}{\alpha}\right)^2}\right) - \frac{1}{2}\left(\frac{\delta^2\gamma_{ND}}{1-\delta\left(\frac{1}{\alpha}\right)^2}\right)^2 \\ =\ &\delta\left(\frac{\delta^4\gamma^2_{ND}}{\left(1-\delta\left(\frac{1}{\alpha}\right)^2\right)^2}\left(\frac{1}{\alpha}\right)^2 - \frac{1}{2}\frac{\delta^4\gamma^2_{ND}}{\left(1-\delta\left(\frac{1}{\alpha}\right)^2\right)^2}\left(\frac{1}{\alpha}\right)^2\right) + \frac{\delta^4\gamma^2_{ND}}{1-\delta\left(\frac{1}{\alpha}\right)^2} - \frac{1}{2}\frac{\delta^4\gamma^2_{ND}}{\left(1-\delta\left(\frac{1}{\alpha}\right)^2\right)^2} \\ =\ &\frac{1}{2}\delta\frac{\delta^4\gamma^2_{ND}}{\left(1-\delta\left(\frac{1}{\alpha}\right)^2\right)^2}\left(\frac{1}{\alpha}\right)^2 + \frac{2\delta^4\gamma^2_{ND}\left(1-\delta\left(\frac{1}{\alpha}\right)^2\right)}{2\left(1-\delta\left(\frac{1}{\alpha}\right)^2\right)^2} - \frac{1}{2}\frac{\delta^4\gamma^2_{ND}}{\left(1-\delta\left(\frac{1}{\alpha}\right)^2\right)^2} \\ =\ &\frac{1}{2}\frac{\delta^4\gamma^2_{ND}}{\left(1-\delta\left(\frac{1}{\alpha}\right)^2\right)^2}\left(\delta\left(\frac{1}{\alpha}\right)^2 + 2\left(1-\delta\left(\frac{1}{\alpha}\right)^2\right) - 1\right) \\ =\ &\frac{1}{2}\frac{\delta^4\gamma^2_{ND}}{\left(1-\delta\left(\frac{1}{\alpha}\right)^2\right)^2}\left(1-\delta\left(\frac{1}{\alpha}\right)^2\right) = \frac{1}{2}\frac{\delta^4\gamma^2_{ND}}{1-\delta\left(\frac{1}{\alpha}\right)^2} \\ \Rightarrow\ &U^*_{R,NP_1^{never}} = W^*_{NP_1^{never}} = \frac{1}{2}e^*_{0,NP_1^{never}}\cdot\delta^2\gamma_{ND}.\end{aligned}$$

7.1.4 Researcher publishes two subsequent papers and never shares (*2-Never*)

Under *2-Never*, R maximizes

$$U_{R,NP_2^{never}} = \delta\left(e_0 e_1 \frac{1}{\alpha} - \frac{1}{2}e_1^2\right) + \delta^2\left(e_0 e_2 \frac{1}{c\cdot\alpha} - \frac{1}{2}e_2^2\right) + \delta^3\gamma_{ND}e_0 - \frac{1}{2}e_0^2.$$

The FOCs are given by:

$$\begin{aligned}
\frac{\partial U_{R,NP_2^{never}}}{\partial e_0} &= 0 \Rightarrow \delta\left(e_1\frac{1}{\alpha}\right) + \delta^2\left(e_2\frac{1}{c\cdot\alpha}\right) + \delta^3\gamma_{ND} = e_0\left(e_1, e_2\right)\\
\frac{\partial U_{R,NP_2^{never}}}{\partial e_1} &= 0 \Rightarrow e_0\frac{1}{\alpha} = e_1\left(e_0\right)\\
\frac{\partial U_{R,NP_2^{never}}}{\partial e_2} &= 0 \Rightarrow e_0\frac{1}{c\cdot\alpha} = e_2\left(e_0\right)
\end{aligned}$$

By plugging $e_1(e_0)$ and $e_2(e_0)$ in $e_0(e_1, e_2)$, we obtain

$$\begin{aligned}
e_0 &= \delta\left(\left(e_0\frac{1}{\alpha}\right)\frac{1}{\alpha}\right) + \delta^2\left(\left(e_0\frac{1}{c\cdot\alpha}\right)\frac{1}{c\cdot\alpha}\right) + \delta^3\gamma_{ND}\\
&\iff e_0 = \delta e_0\left(\frac{1}{\alpha}\right)^2 + \delta^2 e_0\left(\frac{1}{c\cdot\alpha}\right)^2 + \delta^3\gamma_{ND} \iff e_0\left(1 - \delta\left(\frac{1}{\alpha}\right)^2 - \delta^2\left(\frac{1}{c\cdot\alpha}\right)^2\right) = \delta^3\gamma_{ND}\\
&\iff e^*_{0,NP_2^{never}} = \frac{\delta^3\gamma_{ND}}{1 - \delta\left(\frac{1}{\alpha}\right)^2 - \delta^2\left(\frac{1}{c\cdot\alpha}\right)^2}.
\end{aligned}$$

By setting $e^*_{0,NP_2^{never}}$ in $e_1(e_0)$ and $e_2(e_0)$ we obtain:

$$e^*_{1,NP_2^{never}} = e^*_{0,NP_2^{never}}\cdot\frac{1}{\alpha}, \quad e^*_{2,NP_2^{never}} = e^*_{0,NP_2^{never}}\cdot\frac{1}{c\cdot\alpha}.$$

By setting $e_1(e_0)$ and $e_2(e_0)$ in $U_{R,NP_2^{never}}$, we obtain:

$$\begin{aligned}
&\delta\left(e_0\left(e_0\frac{1}{\alpha}\right)\frac{1}{\alpha} - \frac{1}{2}\left(e_0\frac{1}{\alpha}\right)^2\right) + \delta^2\left(e_0\left(e_0\frac{1}{c\cdot\alpha}\right)\frac{1}{c\cdot\alpha} - \frac{1}{2}\left(e_0\frac{1}{c\cdot\alpha}\right)^2\right) + \delta^3\gamma_{ND}e_0 - \frac{1}{2}e_0^2\\
= \;& \delta\left(e_0^2\left(\frac{1}{\alpha}\right)^2 - \frac{1}{2}e_0^2\left(\frac{1}{\alpha}\right)^2\right) + \delta^2\left(e_0^2\left(\frac{1}{c\cdot\alpha}\right)^2 - \frac{1}{2}e_0^2\left(\frac{1}{c\cdot\alpha}\right)^2\right) + \delta^3\gamma_{ND}e_0 - \frac{1}{2}e_0^2\\
= \;& \delta\frac{1}{2}e_0^2\left(\frac{1}{\alpha}\right)^2 + \delta^2\frac{1}{2}e_0^2\left(\frac{1}{c\cdot\alpha}\right)^2 - \frac{1}{2}e_0^2 + \frac{2\delta^3\gamma_{ND}e_0}{2}\\
= \;& \frac{1}{2}\left(e_0^2\left(\delta\left(\frac{1}{\alpha}\right)^2 + \delta^2\left(\frac{1}{c\cdot\alpha}\right)^2 - 1\right) + 2\delta^3\gamma_{ND}e_0\right).
\end{aligned}$$

Now plug in $e^*_{0,NP_2^{never}}$:

$$\begin{aligned}
& \frac{1}{2}\left(-\frac{\delta^6\gamma_{ND}^2}{\left(1-\delta\left(\frac{1}{\alpha}\right)^2-\delta^2\left(\frac{1}{c\cdot\alpha}\right)^2\right)^2}\left(1-\delta\left(\frac{1}{\alpha}\right)^2-\delta^2\left(\frac{1}{c\cdot\alpha}\right)^2\right)+2\delta^3\gamma_{ND}\left(\frac{\delta^3\gamma_{ND}}{1-\delta\left(\frac{1}{\alpha}\right)^2-\delta^2\left(\frac{1}{c\cdot\alpha}\right)^2}\right)\right) \\
= & \frac{1}{2}\left(-\frac{\delta^6\gamma_{ND}^2}{1-\delta\left(\frac{1}{\alpha}\right)^2-\delta^2\left(\frac{1}{c\cdot\alpha}\right)^2}+\frac{2\delta^6\gamma_{ND}^2}{1-\delta\left(\frac{1}{\alpha}\right)^2-\delta^2\left(\frac{1}{c\cdot\alpha}\right)^2}\right) \\
= & U^*_{R,NP_2^{never}}=\frac{1}{2}\frac{\delta^6\gamma_{ND}^2}{1-\delta\left(\frac{1}{\alpha}\right)^2-\delta^2\left(\frac{1}{c\cdot\alpha}\right)^2}=\frac{1}{2}e^*_{0,NP_2^{never}}\cdot\delta^3\gamma_{ND}.
\end{aligned}$$

7.2 Appendix B: Optimal efforts, utility and welfare under FPP

7.2.1 FPP without strategic delay

FPP without strategic delay is equivalent to NP with data sharing at $t=1$.

7.2.2 FPP with strategic delay

Under FPP with strategic delay, R maximizes

$$U_{R,FPP_{SD}}=\delta\left(\delta e_0e_1\frac{1}{\alpha}(1+\kappa)-\frac{1}{2}e_1^2\right)+\delta^2\left(e_0e_2\frac{1}{c\cdot\alpha}(1+\kappa)-\frac{1}{2}e_2^2\right)+\delta^3\gamma_De_0-\frac{1}{2}e_0^2.$$

From the FOCs we obtain

$$\begin{aligned}
\frac{\partial U_{R,FPP_{SD}}}{\partial e_0} &= 0\Rightarrow\delta^2\left(e_1\frac{1}{\alpha}(1+\kappa)+e_2\frac{1}{c\cdot\alpha}(1+\kappa)\right)+\delta^3\gamma_D=e_0\left(e_1,e_2\right), \\
\frac{\partial U_{R,FPP_{SD}}}{\partial e_1} &= 0\Rightarrow\delta e_0\frac{1}{\alpha}(1+\kappa)=e_1\left(e_0\right), \\
\frac{\partial U_{R,FPP_{SD}}}{\partial e_2} &= 0\Rightarrow e_0\frac{1}{c\cdot\alpha}(1+\kappa)=e_2\left(e_0\right).
\end{aligned}$$

By plugging $e_1(e_0)$ and $e_2(e_0)$ in $e_0(e_1, e_2)$, we obtain

$$\begin{aligned}
e_0 &= \delta^2\left(\left(\delta e_0\frac{1}{\alpha}(1+\kappa)\right)\frac{1}{\alpha}(1+\kappa)+\left(e_0\frac{1}{c\cdot\alpha}(1+\kappa)\right)\frac{1}{c\cdot\alpha}(1+\kappa)\right)+\delta^3\gamma_D\\
&\iff e_0=\delta^2\left(\delta e_0\left(\frac{1}{\alpha}\right)^2(1+\kappa)^2+e_0\left(\frac{1}{c\cdot\alpha}\right)^2(1+\kappa)^2\right)+\delta^3\gamma_D\\
&\iff e_0-\delta^2 e_0\left(\left(\delta\left(\frac{1}{\alpha}\right)^2+\left(\frac{1}{c\cdot\alpha}\right)^2\right)(1+\kappa)^2\right)=\delta^3\gamma_D\\
&\iff e^*_{0,FPP_{SD}}=\frac{\delta^3\gamma_D}{1-\delta^3\left(\frac{1}{\alpha}\right)^2(1+\kappa)^2-\delta^2\left(\frac{1}{c\cdot\alpha}\right)^2(1+\kappa)^2}.
\end{aligned}$$

By setting $e^*_{0,FPP_{SD}}$ in $e_1(e_0)$ and $e_2(e_0)$ we obtain:

$$e^*_{1,FPP_{SD}}=\delta\cdot e^*_{0,FPP_{SD}}\cdot\frac{1}{\alpha}(1+\kappa),\quad e^*_{2,FPP_{SD}}=e^*_{0,FPP_{SD}}\cdot\frac{1}{c\cdot\alpha}(1+\kappa).$$

By setting $e_1(e_0)$ and $e_2(e_0)$ in $U_{R,FPP_{SD}}$, we obtain:

$$\begin{aligned}
&\delta\left(\delta e_0\left(\delta e_0\frac{1}{\alpha}(1+\kappa)\right)\frac{1}{\alpha}(1+\kappa)-\frac{1}{2}\left(\delta e_0\frac{1}{\alpha}(1+\kappa)\right)^2\right)\\
&+\delta^2\left(e_0\left(e_0\frac{1}{c\cdot\alpha}(1+\kappa)\right)\frac{1}{c\cdot\alpha}(1+\kappa)-\frac{1}{2}\left(e_0\frac{1}{c\cdot\alpha}(1+\kappa)\right)^2\right)+\delta^3\gamma_D e_0-\frac{1}{2}e_0^2\\
=\;&\delta\left(\delta^2e_0^2\left(\frac{1}{\alpha}\right)^2(1+\kappa)^2-\frac{1}{2}\delta^2e_0^2\left(\frac{1}{\alpha}\right)^2(1+\kappa)^2\right)\\
&+\delta^2\left(e_0^2\left(\frac{1}{c\cdot\alpha}\right)^2(1+\kappa)^2-\frac{1}{2}e_0^2\left(\frac{1}{c\cdot\alpha}\right)^2(1+\kappa)^2\right)+\delta^3\gamma_D e_0-\frac{1}{2}e_0^2\\
=\;&\frac{1}{2}\delta^3e_0^2\left(\frac{1}{\alpha}\right)^2(1+\kappa)^2+\frac{1}{2}\delta^2e_0^2\left(\frac{1}{c\cdot\alpha}\right)^2(1+\kappa)^2+\delta^3\gamma_D e_0-\frac{1}{2}e_0^2\\
=\;&-\frac{1}{2}e_0^2\left(1-\delta^3\left(\frac{1}{\alpha}\right)^2(1+\kappa)^2-\delta^2\left(\frac{1}{c\cdot\alpha}\right)^2(1+\kappa)^2\right)+\delta^3\gamma_D e_0.
\end{aligned}$$

Now plug in $e^*_{0,FPP_{SD}}$:

$$
\begin{aligned}
& -\frac{1}{2}\frac{\delta^6\gamma_D^2\left(1-\delta^3\left(\frac{1}{\alpha}\right)^2(1+\kappa)^2-\delta^2\left(\frac{1}{c\cdot\alpha}\right)^2(1+\kappa)^2\right)}{\left(1-\delta^3\left(\frac{1}{\alpha}\right)^2(1+\kappa)^2-\delta^2\left(\frac{1}{c\cdot\alpha}\right)^2(1+\kappa)^2\right)^2}+\frac{\delta^6\gamma_D^2}{1-\delta^3\left(\frac{1}{\alpha}\right)^2(1+\kappa)^2-\delta^2\left(\frac{1}{c\cdot\alpha}\right)^2(1+\kappa)^2} \\
= \quad & \frac{1}{2}\frac{\delta^6\gamma_D^2}{1-\delta^3\left(\frac{1}{\alpha}\right)^2(1+\kappa)^2-\delta^2\left(\frac{1}{c\cdot\alpha}\right)^2(1+\kappa)^2} \\
\iff \quad & U_{R,FPP_{SD}}=\frac{1}{2}e^*_{0,FPP_{SD}}\delta^3\gamma_D.
\end{aligned}
$$

7.3 Appendix C: Proofs of Propositions

Proof of Proposition 0: Since the denominators of optimal efforts depend negatively on δ, the efforts are positive for all $\delta \in (0,1]$ if all of the following conditions hold:

$$e^*_{0,NP_2^{never}}(\delta=1)>0 \text{ if } 1-\frac{1}{\alpha^2}-\frac{1}{c^2\alpha^2}>0 \tag{1}$$

$$e^*_{0,NP_1^{never}}(\delta=1)>0 \text{ if } 1-\frac{1}{\alpha^2}>0 \tag{2}$$

$$e^*_{0,NP_2^{share}}(\delta=1)>0 \text{ if } 1-\frac{(1+\kappa)^2}{\alpha^2}-\frac{(1+\kappa)^2}{c^2\alpha^2}>0 \tag{3}$$

$$e^*_{0,NP_1^{share}}(\delta=1)>0 \text{ if } 1-\frac{(1+\kappa)^2}{\alpha^2}>0 \tag{4}$$

$$e^*_{0,FPP_{SD}}(\delta=1)>0 \text{ if } 1-\frac{(1+\kappa)^2}{\alpha^2}-\frac{(1+\kappa)^2}{c^2\alpha^2}>0 \tag{5}$$

It is clear that conditions (1), (2) and (4) hold whenever condition (3) or, equivalently, (5) holds. We can rewrite (3) as

$$c^2\alpha^2-(1+c^2)(1+\kappa)^2>0 \Rightarrow (1+\kappa)^2<\frac{c^2}{1+c^2}\alpha^2.$$

Proof of Proposition 1: We first show that for $\xi \geq 1$ it always holds that $U_{R,NP_1^{share}} \geq U_{R,NP_1^{never}}$ and $U_{R,NP_2^{share}} \geq U_{R,NP_2^{never}}$.

The first inequality is

$$\frac{1}{2}\frac{\delta^4\gamma_D^2}{1-\delta\left(\frac{1}{\alpha}\right)^2(1+\kappa)^2}\geq\frac{1}{2}\frac{\delta^4\gamma_{ND}^2}{1-\delta\left(\frac{1}{\alpha}\right)^2} \Leftrightarrow \left(\frac{\gamma_D}{\gamma_{ND}}\right)^2\geq\frac{1-\delta\left(\frac{1}{\alpha}\right)^2(1+\kappa)^2}{1-\delta\left(\frac{1}{\alpha}\right)^2}, \tag{6}$$

which holds whenever $\xi=\frac{\gamma_D}{\gamma_{ND}}\geq 1$.

The second inequality is

$$\frac{1}{2}\frac{\delta^6\gamma_D^2}{1-\delta\left(\frac{1}{\alpha}\right)^2(1+\delta\kappa)^2-\delta^2\left(\frac{1}{c\alpha}\right)^2(1+\kappa)^2}\geq\frac{1}{2}\frac{\delta^6\gamma_{ND}^2}{1-\delta\left(\frac{1}{\alpha}\right)^2-\delta^2\left(\frac{1}{c\alpha}\right)^2}\Leftrightarrow$$
$$\left(\frac{\gamma_D}{\gamma_{ND}}\right)^2\geq 1\geq\frac{1-\delta\left(\frac{1}{\alpha}\right)^2(1+\delta\kappa)^2-\delta^2\left(\frac{1}{c\alpha}\right)^2(1+\kappa)^2}{1-\delta\left(\frac{1}{\alpha}\right)^2-\delta^2\left(\frac{1}{c\alpha}\right)^2},\tag{7}$$

which holds whenever $\xi=\frac{\gamma_D}{\gamma_{ND}}\geq 1$, since $(1+\kappa)^2,(1+\delta\kappa)^2\geq 1$.

(i) We define

$$\xi_l(\kappa):=\min_{\delta\in(0,1]}\left\{\sqrt{\frac{1-\frac{\delta}{\alpha^2}(1+\kappa)^2}{1-\frac{\delta}{\alpha^2}}},\ \sqrt{\frac{1-\frac{\delta}{\alpha^2}(1+\delta\kappa)^2-\frac{\delta^2}{\alpha^2c^2}(1+\kappa)^2}{1-\frac{\delta}{\alpha^2}-\frac{\delta^2}{\alpha^2c^2}}}\right\}.\tag{8}$$

Obviously, for $\xi\leq\xi_l(\kappa)$ the Researcher will choose to never share and $\xi_l(\kappa)$ negatively depends on κ. $\xi_l(\kappa)$ is positive for positive efforts. We now simplify (8).

Since $\frac{(1+\kappa)^2}{\alpha^2}>\frac{1}{\alpha^2}$, $\sqrt{\frac{1-\frac{\delta}{\alpha^2}(1+\kappa)^2}{1-\frac{\delta}{\alpha^2}}}$ is monotonically decreasing in δ and

$$\min_{\delta\in(0,1]}\sqrt{\frac{1-\frac{\delta}{\alpha^2}(1+\kappa)^2}{1-\frac{\delta}{\alpha^2}}}=\sqrt{\frac{1-\frac{(1+\kappa)^2}{\alpha^2}}{1-\frac{1}{\alpha^2}}}.\tag{9}$$

Since $\frac{(1+\delta\kappa)^2}{\alpha^2}>\frac{1}{\alpha^2}$ and $\frac{(1+\kappa)^2}{c^2\alpha^2}>\frac{1}{c^2\alpha^2}$, $\sqrt{\frac{1-\frac{\delta}{\alpha^2}(1+\delta\kappa)^2-\frac{\delta^2}{\alpha^2c^2}(1+\kappa)^2}{1-\frac{\delta}{\alpha^2}-\frac{\delta^2}{\alpha^2c^2}}}$ is monotonically decreasing in δ and

$$\min_{\delta\in(0,1]}\sqrt{\frac{1-\frac{\delta}{\alpha^2}(1+\delta\kappa)^2-\frac{\delta^2}{\alpha^2c^2}(1+\kappa)^2}{1-\frac{\delta}{\alpha^2}-\frac{\delta^2}{\alpha^2c^2}}}=\sqrt{\frac{1-\frac{(1+\kappa)^2}{\alpha^2}-\frac{(1+\kappa)^2}{\alpha^2c^2}}{1-\frac{1}{\alpha^2}-\frac{1}{\alpha^2c^2}}}.\tag{10}$$

Finally,

$$\sqrt{\frac{1-\frac{(1+\kappa)^2}{\alpha^2}-\frac{(1+\kappa)^2}{\alpha^2c^2}}{1-\frac{1}{\alpha^2}-\frac{1}{\alpha^2c^2}}}<\sqrt{\frac{1-\frac{(1+\kappa)^2}{\alpha^2}}{1-\frac{1}{\alpha^2}}}$$
$$\Leftrightarrow\left(1-\frac{(1+\kappa)^2}{\alpha^2}\right)\left(1-\frac{1}{\alpha^2}\right)-\frac{(1+\kappa)^2}{\alpha^2c^2}\left(1-\frac{1}{\alpha^2}\right)$$
$$<\left(1-\frac{(1+\kappa)^2}{\alpha^2}\right)\left(1-\frac{1}{\alpha^2}\right)-\frac{1}{\alpha^2c^2}\left(1-\frac{(1+\kappa)^2}{\alpha^2}\right)$$
$$\Leftrightarrow\frac{1}{\alpha^2c^2}\left(1-\frac{(1+\kappa)^2}{\alpha^2}\right)<\frac{(1+\kappa)^2}{\alpha^2c^2}\left(1-\frac{1}{\alpha^2}\right)$$
$$\Leftrightarrow\alpha^2-(1+\kappa)^2<\alpha^2(1+\kappa)^2-(1+\kappa)^2\Leftrightarrow 1<(1+\kappa)^2\Leftrightarrow 0<\kappa,\tag{11}$$

so we can state

$$\xi_l = \sqrt{\frac{1 - \frac{(1+\kappa)^2}{\alpha^2} - \frac{(1+\kappa)^2}{\alpha^2 c^2}}{1 - \frac{1}{\alpha^2} - \frac{1}{\alpha^2 c^2}}}. \tag{12}$$

(ii) Now we aim to show that for $\xi_l < \xi < 1$ there always exists some $\delta_h(\kappa, \xi) > 0$, such that for every $\delta \leq \delta_h(\kappa, \xi)$ it holds $U_{R,NP_1^{never}} \geq U_{R,NP_1^{share}}$, $U_{R,NP_1^{never}} \geq U_{R,NP_2^{never}}$ and $U_{R,NP_1^{never}} \geq U_{R,NP_2^{share}}$.

So, the first inequality is equivalent to

$$\begin{aligned} \frac{1}{2}\frac{\delta^4\gamma_D^2}{1 - \delta\left(\frac{1}{\alpha}\right)^2(1+\kappa)^2} \leq \frac{1}{2}\frac{\delta^4\gamma_{ND}^2}{1 - \delta\left(\frac{1}{\alpha}\right)^2} \;\Leftrightarrow \xi^2 - \frac{\delta}{\alpha^2}\xi^2 \leq 1 - \frac{\delta}{\alpha^2}(1+\kappa)^2 \;\Leftrightarrow \\ \delta \leq \frac{\alpha^2(1-\xi^2)}{(1+\kappa)^2 - \xi^2}, \end{aligned} \tag{13}$$

where the expression on the right-hand side (r.h.s.) is always positive for $\xi < 1$ and $\kappa > 0$. Thus, there exists some $\delta_{0_0}(\kappa, \xi) := \frac{\alpha^2(1-\xi^2)}{(1+\kappa)^2-\xi^2} \in [0, 1]$ such that $U_{R,NP_1^{never}} \geq U_{R,NP_1^{share}}$ for all $\delta \leq \delta_{0_0}$.

The second inequality is equivalent to

$$\begin{aligned} \frac{1}{2}\frac{\delta^4\gamma_{ND}^2}{1 - \frac{\delta}{\alpha^2}} \geq \frac{1}{2}\frac{\delta^6\gamma_{ND}^2}{1 - \frac{\delta}{\alpha^2} - \frac{\delta^2}{\alpha^2 c^2}} \;\Leftrightarrow 1 - \frac{\delta}{\alpha^2} - \frac{\delta^2}{\alpha^2 c^2} \geq \delta^2 - \frac{\delta^3}{\alpha^2} \;\Leftrightarrow \\ c^2\delta^3 - (1 + \alpha^2 c^2)\delta^2 - c^2\delta + \alpha^2 c^2 \geq 0. \end{aligned} \tag{14}$$

Denoting the function on the left-hand side (l.h.s.) $lhs(\delta)$, we compute $lhs(0) = \alpha^2 c^2 > 0$, $lhs(1) = -1 < 0$ and

$$\frac{\partial lhs\,(\delta)}{\partial \delta} = 3c^2\delta^2 - 2(1 + \alpha^2 c^2)\delta - c^2 < 0 \tag{15}$$

for all $\delta \in [0, 1]$. So there exists some $\delta_{0_1} = \delta_{0_1}(\kappa, \xi) \in [0, 1]$ such that $lhs(\delta_{0_1}) = 0$ and $lhs(\delta) > 0$ for all $\delta < \delta_{0_1}$.

The third inequality is equivalent to

$$\begin{aligned} \frac{1}{2}\frac{\delta^4\gamma_{ND}^2}{1 - \frac{\delta}{\alpha^2}} \geq \frac{1}{2}\frac{\delta^6\gamma_D^2}{1 - \frac{\delta}{\alpha^2}(1+\delta\kappa)^2 - \frac{\delta^2}{\alpha^2 c^2}(1+\kappa)^2} \;\Leftrightarrow \\ 1 - \frac{\delta}{\alpha^2} - \frac{2\delta^2}{\alpha^2}\kappa - \frac{\delta^3}{\alpha^2}\kappa^2 - \frac{\delta^2}{\alpha^2 c^2}(1+\kappa)^2 \geq \delta^2\xi^2 - \frac{\delta^3}{\alpha^2}\xi^2 \;\Leftrightarrow \\ \frac{\delta^3}{\alpha^2}(\xi^2 - \kappa^2) - \delta^2\left(\frac{2\kappa}{\alpha^2} + \xi^2 + \frac{(1+\kappa)^2}{\alpha^2 c^2}\right) - \frac{\delta}{\alpha^2} + 1 \geq 0. \end{aligned} \tag{16}$$

Denoting the function on the l.h.s. $LHS(\delta)$, we find that $LHS(0) = 1 > 0$ and

$$LHS(1) = (1 - \xi^2)\left(1 - \frac{1}{\alpha^2}\right) - \frac{2\kappa}{\alpha^2} - \frac{\kappa^2}{\alpha^2} - \frac{(1+\kappa)^2}{\alpha^2 c^2}. \tag{17}$$

Further,

$$\frac{\partial LHS\,(\delta)}{\partial \delta} = 3\frac{\delta^2}{\alpha^2}(\xi^2 - \kappa^2) - 2\delta\left(\frac{2\kappa}{\alpha^2} + \xi^2 + \frac{(1+\kappa)^2}{\alpha^2 c^2}\right) - \frac{1}{\alpha^2} < 0 \tag{18}$$

for all $\delta \in [0, 1]$.

If $LHS(1) < 0$, then there exists some $\delta_{0_2} = \delta_{0_2}(\kappa, \xi)$ such that $LHS(\delta_{0_2}) = 0$ and $LHS(\delta) > 0$ for all $\delta < \delta_{0_2}$. For completeness, define $\delta_{0_2} \equiv 1$ in case $LHS(1) > 0$.

Combining our considerations, we define

$$\delta_h(\kappa, \xi) := \min\left\{\delta_{0_0}, \delta_{0_1}, \delta_{0_2}\right\}, \tag{19}$$

which is always positive for $0 \leq \kappa \leq 0.5$ and $\xi_l < \xi < 1$. Clearly, for $\delta < \delta_h$ the Researcher will choose not to share. It remains to show that $\delta_h(\kappa, \xi)$ depends nonpositively on both κ and ξ.

Computing the κ and ξ derivatives of the first term on the r.h.s. of (19), one can see that both of them are negative everywhere. If $\delta_{0_2} = 1$, then the statement is proved, since the first term is the only one that depends on κ and ξ.

We now consider $LHS(\delta)$ for a fixed δ as the function of κ and ξ, denoting it $LHS^\delta(\kappa, \xi)$. Notice that

$$\frac{\partial LHS^\delta(\kappa, \xi)}{\partial \kappa} = -\frac{2\delta^3}{\alpha^2}\kappa - \delta^2\left(\frac{2}{\alpha^2} + \frac{2(1+\kappa)}{\alpha^2 c^2}\right) < 0 \tag{20}$$

and

$$\frac{\partial LHS^\delta(\kappa, \xi)}{\partial \xi} = \frac{2\delta^3}{\alpha^2}\xi - 2\delta^2\xi = 2\delta^2\xi\left(\frac{\delta}{\alpha^2} - 1\right) < 0 \tag{21}$$

for $\delta \in [0, 1]$.

So the function $LHS(\delta)$ is pointwise monotone decreasing separately in κ and ξ. Since $LHS(\delta)$ is monotonically decreasing in point δ_{0_2}, it follows that $\delta_{0_2}(\kappa, \xi)$ depends nonpositively on both κ and ξ.

Proof of Proposition 2: See p. 33 for the proofs of Proposition 2.

Proof of Proposition 3: We prove the statement generally for all $0 \leq \kappa \leq 0.5$ and $c, \alpha > 1$.

The Researcher chooses to delay strategically under FPP if and only if $U_{R,FPP_{NSD}} \leq U_{R,FPP_{SD}}$, i.e.

$$\frac{\delta^4\gamma_D^2}{2\left(1-\frac{\delta(1+\kappa)^2}{\alpha^2}\right)} \leq \frac{\delta^6\gamma_D^2}{2\left(1-\frac{\delta^3(1+\kappa)^2}{\alpha^2}-\frac{\delta^2(1+\kappa)^2}{c^2\alpha^2}\right)} \Leftrightarrow 1-\frac{\delta^3(1+\kappa)^2}{\alpha^2}-\frac{\delta^2(1+\kappa)^2}{c^2\alpha^2} \leq \delta^2-\frac{\delta^3(1+\kappa)^2}{\alpha^2}$$

$$\Leftrightarrow 1 \leq \delta^2+\frac{\delta^2(1+\kappa)^2}{c^2\alpha^2} \Leftrightarrow \sqrt{\frac{1}{1+\frac{(1+\kappa)^2}{c^2\alpha^2}}} \leq \delta, \tag{22}$$

where the first step is correct under the assumption of positive efforts. So there is some $\delta_{h_{SD}} = \delta_{h_{SD}}(\kappa) \in [0,1]$ with $U_{R,FPP_{NSD}} \leq U_{R,NP_{SD}}$ for every $\delta \geq \delta_{h_{SD}}$. Note that $\delta_{h_{SD}}(\kappa)$ depends negatively on κ.

Proof of Proposition 4: (i) We define δ_{h_1} so that $U_{NP_1^{share}} \geq U_{NP_2^{share}}$ for $\delta \leq \delta_{h_1}$. For this consider

$$\frac{1}{2}\frac{\delta^4\gamma_D^2}{1-\frac{\delta}{\alpha^2}(1+\kappa)^2} \geq \frac{1}{2}\frac{\delta^6\gamma_D^2}{1-\frac{\delta}{\alpha^2}(1+\delta\kappa)^2-\frac{\delta^2}{c^2\alpha^2}(1+\kappa)^2} \Leftrightarrow$$

$$1-\frac{\delta}{\alpha^2}-\delta^2\frac{2\kappa}{\alpha^2}-\delta^3\frac{\kappa^2}{\alpha^2}-\delta^2\frac{(1+\kappa)^2}{c^2\alpha^2} \geq \delta^2-\frac{\delta^3}{\alpha^2}(1+\kappa)^2 \Leftrightarrow$$

$$\delta^3\frac{1+2\kappa}{\alpha^2}-\delta^2\left[1+\frac{2\kappa}{\alpha^2}+\frac{(1+\kappa)^2}{c^2\alpha^2}\right]-\frac{\delta}{\alpha^2}+1 \geq 0. \tag{23}$$

Denoting l.h.s. of the last expression $LHS(\delta)$ we compute $LHS(0) = 1 > 0$ and

$$LHS(1) = \frac{1+2\kappa}{\alpha^2}-1-\frac{2\kappa}{\alpha^2}-\frac{(1+\kappa)^2}{c^2\alpha^2}-\frac{1}{\alpha^2}+1 = -\frac{(1+\kappa)^2}{c^2\alpha^2} < 0. \tag{24}$$

The first derivative of LHS with respect to δ is

$$LHS'_\delta = 3\delta^2\frac{1+2\kappa}{\alpha^2}-2\delta\left[1+\frac{2\kappa}{\alpha^2}+\frac{(1+\kappa)^2}{c^2\alpha^2}\right]-\frac{1}{\alpha^2}. \tag{25}$$

We compute $LHS'_\delta(0) = -\frac{1}{\alpha^2} < 0$ and

$$LHS'_\delta(1) = \frac{2(1+\kappa)}{\alpha^2}-2-2\frac{(1+\kappa)^2}{c^2\alpha^2} < 0, \tag{26}$$

because positive efforts imply $\frac{1+\kappa}{\alpha^2} < 1$ (see (4)). The derivative of LHS'_δ with respect to δ is monotone as the derivative of any quadratic function and we conclude that $LHS(\delta)$ is monotonically decreasing for all $\delta \in [0,1]$. Accordingly, $\delta_{h_1} \in (0,1)$ is well-defined as the

first positive root of $LHS(\delta) = 0$, i.e. for $\delta \leq \delta_{h_1}$ it indeed holds $U_{NP_1^{share}} \geq U_{NP_2^{share}}$.

Now by the proof of Proposition 1, for $\xi \geq 1$ and $\delta \leq \delta_{h_1}$ the Researcher will choose to share after one publication under NP. Whenever the Researcher chooses to share after one publication under NP, she also chooses not to delay strategically under FPP. We show this indirectly, proving that sharing after 2 publications strictly dominates strategic delay. This would imply that the Researcher will not choose strategic delay, because she does not choose sharing even after 2 publications. $U_{NP_2^{share}} \geq U_{FPP_{SD}}$ is equivalent to

$$\begin{gathered}
\frac{1}{2}\frac{\delta^6\gamma_D^2}{1-\frac{\delta}{\alpha^2}(1+\delta\kappa)^2-\frac{\delta^2}{c^2\alpha^2}(1+\kappa)^2} \geq \frac{1}{2}\frac{\delta^6\gamma_D^2}{\left(1-\frac{\delta^3(1+\kappa)^2}{\alpha^2}-\frac{\delta^2(1+\kappa)^2}{c^2\alpha^2}\right)} \Leftrightarrow \\
1-\frac{\delta^3(1+\kappa)^2}{\alpha^2}-\frac{\delta^2(1+\kappa)^2}{c^2\alpha^2} \geq 1-\frac{\delta}{\alpha^2}(1+\delta\kappa)^2-\frac{\delta^2}{c^2\alpha^2}(1+\kappa)^2 \Leftrightarrow \\
-\frac{\delta^3(1+\kappa)^2}{\alpha^2} \geq -\frac{\delta}{\alpha^2}(1+\delta\kappa)^2 \Leftrightarrow (1+\delta\kappa)^2 \geq \delta^2(1+\kappa)^2 \Leftrightarrow \\
1+2\delta\kappa+\delta^2\kappa^2 \geq \delta^2+2\delta^2\kappa+\delta^2\kappa^2 \Leftrightarrow 1+2\delta\kappa \geq \delta^2+2\delta^2\kappa \Leftrightarrow \\
1-\delta^2\left(1+2\kappa\right)+2\delta\kappa \geq 0, \qquad (27)
\end{gathered}$$

which is true for all $\delta \in (0, 1]$.

Since sharing after one publication is the same as not delaying strategically, it is clear that for $\xi \geq 1$ and $\delta \leq \delta_{h_1}$ and, analogously, for those cases where $\xi < 1$ and the Researcher chooses to share after one publication under NP, transition to FPP does not influence welfare. We first show the dependence of $\delta_{h_1}(\kappa)$ on κ and then discuss the case $\xi < 1$.

Consider $LHS(\delta)$ for a fixed δ as the function of κ and ξ, denoting it $LHS^{\delta}(\kappa, \xi)$. Notice that

$$\frac{\partial LHS^{\delta}(\kappa,\xi)}{\partial\kappa} = \delta^3\frac{2}{\alpha^2} - \delta^2\left[\frac{2}{\alpha^2}+\frac{2}{c^2\alpha^2}+\frac{2\kappa}{c^2\alpha^2}\right] < 0. \qquad (28)$$

for $\delta \in [0, 1]$.

So the function $LHS(\delta)$ is pointwise monotone decreasing in κ. Since $LHS(\delta)$ is monotonically decreasing in point δ_{h_1}, it follows that $\delta_{h_1}(\kappa)$ depends negatively on κ.

For the case $\xi < 1$ we aim to find conditions on δ and ξ under which the utility of sharing after one publication is higher than other utilities. First, we already know from the beginning of this proof that $U_{R,NP_1^{share}} \geq U_{R,NP_2^{share}}$ for $\delta \leq \delta_{h_1}$. Second, we know from the proof of Proposition 1 that for

$$\xi \geq \xi_{h_1} := \sqrt{\frac{1-\frac{\delta}{\alpha^2}(1+\delta\kappa)^2-\frac{\delta^2}{c^2\alpha^2}(1+\kappa)^2}{1-\frac{\delta}{\alpha^2}-\frac{\delta^2}{c^2\alpha^2}}} \qquad (29)$$

it holds that $U_{R,NP_2^{share}} \geq U_{R,NP_2^{never}}$. So for $\xi_{h_1} \leq \xi < 1$ and $\delta \leq \delta_{h_1}$ the Researcher chooses to make one publication and then either to share or not to share. As we know from the proof of Proposition 1(ii) (see (13)), this last choice is met in favour of sharing in case

$$\delta \geq \delta_{0_0}(\xi) := \frac{\alpha^2(1-\xi^2)}{(1+\kappa)^2-\xi^2}. \tag{30}$$

Now $\delta_{0_0}(1) = 0$ and

$$\begin{aligned}\delta'_{0_0}(\xi) &= \frac{-2\alpha^2\xi((1+\kappa)^2-\xi^2)+2\alpha^2\xi(1-\xi^2)}{((1+\kappa)^2-\xi^2)^2} \\ &= \frac{2\alpha^2\xi(1-(1+\kappa)^2)}{((1+\kappa)^2-\xi^2)^2} < 0\end{aligned} \tag{31}$$

for $\xi > 0$. Since δ_{h_1} does not depend on ξ, there exists some ξ_{h_2} such that $\delta_{0_0} \leq \delta_{h_1}$ for every $\xi \geq \xi_{h_2}$.

Defining $\xi_h := \max\{\xi_{h_1},\ \xi_{h_2}\}$ we get that the Researcher chooses to share after one publication if $\xi_h \leq \xi \leq 1$ and $\delta_{0_0} \leq \delta \leq \delta_{h_1}$. Under these conditions $\delta_{0_0} = \delta_h$, since the relevant term in the definition of δ_h is exactly δ_{0_0}. We get the statement.

(ii) (1) By Proposition 3 and the proof of Proposition 1 (ii), we know, that if $\xi_l < \xi \leq 1$ and $\delta < \delta_{h_0} := \min\{\delta_h, \delta_{h_{SD}}\}$, the Researcher will choose to publish one paper and never share under NP or not to delay strategically under FPP. We also already know the dependence of $\xi_l(\kappa)$ and $\delta_{h_0}(\kappa,\xi)$ on κ and ξ. We now show the existence of δ_l and ξ_{l_w} such that for $\xi_{l_w} < \xi \leq 1$ and $\delta_l < \delta < \delta_{h_0}$ it holds that $W_{FPP_{NSD}} > W_{NP_1^{never}}$.

The last expression is equivalent to

$$\begin{aligned}&\frac{\delta^4\gamma_D^2}{2\left(1-\frac{\delta(1+\kappa)^2}{\alpha^2}\right)} + \frac{\delta^6\gamma_D^2}{2c^2\left(1-\frac{\delta(1+\kappa)^2}{\alpha^2}\right)^2} > \frac{\delta^4\gamma_{ND}^2}{2\left(1-\frac{\delta}{\alpha^2}\right)} \quad \Leftrightarrow \\ &\xi^2\left[\frac{1}{\left(1-\frac{\delta(1+\kappa)^2}{\alpha^2}\right)} + \frac{\delta^2}{c^2\left(1-\frac{\delta(1+\kappa)^2}{\alpha^2}\right)^2}\right] > \frac{1}{\left(1-\frac{\delta}{\alpha^2}\right)} \quad \Leftrightarrow \\ &\xi^2 > \left[\frac{1-\frac{\delta}{\alpha^2}}{\left(1-\frac{\delta(1+\kappa)^2}{\alpha^2}\right)} + \frac{\delta^2}{c^2}\frac{1-\frac{\delta}{\alpha^2}}{\left(1-\frac{\delta(1+\kappa)^2}{\alpha^2}\right)^2}\right]^{-1}\end{aligned} \tag{32}$$

Denoting the r.h.s. of the last expression $RHS(\delta)$, we compute $RHS(0) = 1$ and

$$RHS(1) = \left[\frac{1 - \frac{1}{\alpha^2}}{\left(1 - \frac{(1+\kappa)^2}{\alpha^2}\right)} + \frac{1}{c^2} \frac{1 - \frac{1}{\alpha^2}}{\left(1 - \frac{(1+\kappa)^2}{\alpha^2}\right)^2} \right]^{-1} > 0. \tag{33}$$

Moreover, since $\frac{(1+\kappa)^2}{\alpha^2} > \frac{1}{\alpha^2}$, both

$$\frac{1 - \frac{\delta}{\alpha^2}}{\left(1 - \frac{\delta(1+\kappa)^2}{\alpha^2}\right)} \quad \text{and} \quad \frac{1 - \frac{\delta}{\alpha^2}}{\left(1 - \frac{\delta(1+\kappa)^2}{\alpha^2}\right)^2} \tag{34}$$

are monotone increasing and since δ^2 is also monotone increasing, $RHS(\delta)$ is monotonically decreasing in δ.

We conclude that for all ξ_f with $RHS(1) < \xi_f^2 \leq 1$ there exists δ_l such that $W_{FPP_{NSD}} > W_{NP_1^{never}}$ holds for every $\xi_f < \xi < 1$ and $\delta_l < \delta \leq 1$.

Define $\xi_{l_w} := \max\{\sqrt{RHS(1)}, \xi_l\}$. For $\xi = 1$ it holds $\xi^2 \geq RHS(\delta)$ for all $0 < \delta \leq 1$, so the set of points (ξ, δ) with $\xi_{l_w} < \xi \leq 1$ and $\delta_l < \delta < \delta_{h_0}$ is always nonempty.

Finally, from Proposition 1 $\xi_l(\kappa)$ depends negatively on κ and it is easy to see that the same holds for $RHS(1)$.

(2) There might be a case where $\delta_{0_1} < \delta < \min\{\delta_{h_{SD}}, \delta_{0_0}\}$ and $\xi < \xi_{h_1}$, i.e. the Researcher chooses to never share after two publications under NP or not to delay strategically under FPP. Therefore, we consider $W_{FPP_{NSD}} > W_{NP_2^{never}}$. From this, we obtain

$$\xi > \sqrt{\frac{\frac{\delta^2}{1 - \frac{\delta}{\alpha^2} - \frac{\delta^2}{c^2\alpha^2}}}{\left(\frac{1}{1 - \frac{\delta(1+\kappa)^2}{\alpha^2}} + \frac{\delta^2}{c^2\left(1 - \frac{\delta(1+\kappa)^2}{\alpha^2}\right)^2}\right)}}. \tag{35}$$

Denoting the term on the r.h.s. $\xi_{l_3}(\delta, \kappa)$, we obtain that FPP increases welfare if $\delta_{0_1} < \delta < \min\{\delta_{h_{SD}}, \delta_{0_0}\}$ and $\xi_{l_3} < \xi < \xi_{h_1}$. It is easy to see that $\xi_{l_3}(\kappa)$ depends negatively on κ.

Finally, we provide a numerical example which suggests that $W_{FPP_{NSD}} > W_{NP_2^{share}}$ if $\delta_{h_1} < \delta < \delta_{h_{SD}}$. For $c = \alpha = 1.5$ and $\kappa = 0.2$ (see Figure 2), it is straightforward to see that $W_{FPP_{NSD}} > W_{NP_2^{share}}$ if $\delta_{h_1} = 0.8171 < \delta < \delta_{h_{SD}} = 0.8824$.

(iii) As we have seen in part (i) of this proof, sharing after two publications strictly dominates strategic delay. This means that whenever the Researcher chooses to delay strategically under FPP, she could have been better off under NP (even if she would not choose sharing after two publications in this case, as some other NP option could in its turn dominate

sharing after two publications). Since $W_{NP_2^{share}} = U_{NP_2^{share}}$ and $W_{FPP_{SD}} = U_{FPP_{SD}}$, this means that FPP reduces welfare if the Researcher chooses to delay strategically under it, i.e. if $\delta > \delta_{h_{SD}}$.

Now recall from the proof of Proposition 1, (ii), that there exists some δ_h such that for $\delta < \delta_h$ *1-Never* is the dominant strategy under NP. It follows that for $\delta < \min\{\delta_h, \delta_{h_{SD}}\}$ the Researcher chooses to never share after one publication under NP or not to delay strategically under FPP. From (32) we find out that if

$$\xi < \xi_{l_0} := \sqrt{\left[\frac{1-\frac{\delta}{\alpha^2}}{\left(1-\frac{\delta(1+\kappa)^2}{\alpha^2}\right)} + \frac{\delta^2}{c^2}\frac{1-\frac{\delta}{\alpha^2}}{\left(1-\frac{\delta(1+\kappa)^2}{\alpha^2}\right)^2}\right]^{-1}}, \tag{36}$$

then $W_{FPP_{NSD}} < W_{NP_1^{never}}$. That is, for $\xi < \xi_{l_0}$ and $\delta < \min\{\delta_h, \delta_{h_{SD}}\}$ transition to FPP reduces welfare. $\xi_{l_0}(\kappa)$ depends nonpositively on κ.

There may remain a case where $\delta_{0_1} < \delta < \min\{\delta_{h_{SD}}, \delta_{0_0}\}$ and $\xi < \xi_{h_1}$, i.e. the Researcher chooses to never share after two publications under NP or not to delay strategically under FPP. We consider the inequality $W_{FPP_{NSD}} < W_{NP_2^{never}}$ by taking $\xi_{l_3}(\kappa)$ from (35) into account.

Together we get that FPP is welfare reducing when

$$\xi < \xi_{l_1}(\delta) := \begin{cases} \xi_{l_0}(\delta), & \delta < \min\{\delta_h, \delta_{h_{SD}}\} \\ \xi_{l_3}(\delta), & \delta_{0_1} < \delta < \min\{\delta_{h_{SD}}, \delta_{0_0}\} \text{ and } \xi < \xi_{h_1} \end{cases} \tag{37}$$

.

Proof of Proposition 2: (i) (1) We know from the proof of Proposition 1 that for $\xi < \xi_l$ the Researcher chooses to never share and decides to publish two papers if and only if $\delta > \delta_{0_1}$. From Proposition 3, transition to FPP urges the Researcher to delay strategically if and only if $\delta > \delta_{h_{SD}}$. We now compare the efforts to create a data set for different combinations of Researcher's decisions.

$e^*_{0,NP_1^{never}} > e^*_{0,FPP_{NSD}}$ is equivalent to

$$\frac{\delta^2}{1-\frac{\delta}{\alpha^2}} > \frac{\delta^2\xi}{1-\frac{\delta(1+\kappa)^2}{\alpha^2}} \Leftrightarrow \frac{1-\frac{\delta(1+\kappa)^2}{\alpha^2}}{1-\frac{\delta}{\alpha^2}} > \xi. \tag{38}$$

$e^*_{0,NP_2^{never}} > e^*_{0,FPP_{SD}}$ is equivalent to

$$\frac{\delta^3}{1-\frac{\delta}{\alpha^2}-\frac{\delta^2}{c^2\alpha^2}} > \frac{\delta^3\xi}{1-\frac{\delta^3(1+\kappa)^2}{\alpha^2}-\frac{\delta^2(1+\kappa)^2}{c^2\alpha^2}} \Leftrightarrow \frac{1-\frac{\delta^3(1+\kappa)^2}{\alpha^2}-\frac{\delta^2(1+\kappa)^2}{c^2\alpha^2}}{1-\frac{\delta}{\alpha^2}-\frac{\delta^2}{c^2\alpha^2}} > \xi. \tag{39}$$

We define

$$\xi_{l_{e_0}}(\kappa) := \min\left\{\min_{\delta<\min\{\delta_{0_1},\delta_{h_{SD}}\}} \frac{1-\frac{\delta}{\alpha^2}(1+\kappa)^2}{1-\frac{\delta}{\alpha^2}}, \min_{\delta>\max\{\delta_{0_1},\delta_{h_{SD}}\}} \frac{1-\frac{\delta^3}{\alpha^2}(1+\kappa)^2-\frac{\delta^2}{\alpha^2c^2}(1+\kappa)^2}{1-\frac{\delta}{\alpha^2}-\frac{\delta^2}{\alpha^2c^2}}\right\} \tag{40}$$

and notice that in general $\xi_{l_{e_0}} < \xi_l$, since a square root of a number that is smaller than one is bigger than the number itself (there is still a little difference between the second term in the definition of $\xi_{l_{e_0}}$ and the term under the second root in the definition of ξ_l, so we will not use this fact in the proof).

It remains to consider the small interval of δ, with $\delta_{0_1} > \delta > \delta_{h_{SD}}$ or $\delta_{0_1} < \delta < \delta_{h_{SD}}$. $e^*_{0,NP_1^{never}} > e^*_{0,FPP_{SD}}$ is equivalent to

$$\frac{\delta^2}{1-\frac{\delta}{\alpha^2}} > \frac{\delta^3\xi}{1-\frac{\delta^3(1+\kappa)^2}{\alpha^2}-\frac{\delta^2(1+\kappa)^2}{c^2\alpha^2}} \Leftrightarrow \frac{1-\frac{\delta^3(1+\kappa)^2}{\alpha^2}-\frac{\delta^2(1+\kappa)^2}{c^2\alpha^2}}{\delta-\frac{\delta^2}{\alpha^2}} > \xi. \tag{41}$$

$e^*_{0,NP_2^{never}} > e^*_{0,FPP_{NSD}}$ is equivalent to

$$\frac{\delta^3}{1-\frac{\delta}{\alpha^2}-\frac{\delta^2}{c^2\alpha^2}} > \frac{\delta^2\xi}{1-\frac{\delta(1+\kappa)^2}{\alpha^2}} \Leftrightarrow \frac{\delta-\frac{\delta^2(1+\kappa)^2}{\alpha^2}}{1-\frac{\delta}{\alpha^2}-\frac{\delta^2}{c^2\alpha^2}} > \xi. \tag{42}$$

Combining all these results we get that for

$$\xi < \xi_{l_e} := \min\left\{\xi_l, \xi_{l_{e_0}}, \min_{\delta_{0_1}>\delta>\delta_{h_{SD}}} \frac{1-\frac{\delta^3(1+\kappa)^2}{\alpha^2}-\frac{\delta^2(1+\kappa)^2}{c^2\alpha^2}}{\delta-\frac{\delta^2}{\alpha^2}}, \min_{\delta_{0_1}<\delta<\delta_{h_{SD}}} \frac{\delta-\frac{\delta^2(1+\kappa)^2}{\alpha^2}}{1-\frac{\delta}{\alpha^2}-\frac{\delta^2}{c^2\alpha^2}}\right\} \tag{43}$$

transition to FPP reduces efforts to create data set for all $\delta \in [0,1]$. It is easy to see that $\xi_{l_e}(\kappa)$ depends negatively on κ.

(2) When $\xi_{l_e} < \xi < 1$, the Researcher chooses to never share after one publication under NP or not to delay strategically under FPP if and only if $\delta < \min\{\delta_h(\xi), \delta_{h_{SD}}\}$, as we know from the proof of Proposition 1 and from Proposition 3. From (38) we can conclude that in this case FPP reduces efforts to create data set if and only if

$$1-\frac{\delta(1+\kappa)^2}{\alpha^2} > \xi - \frac{\delta}{\alpha^2}\xi \Leftrightarrow \frac{\alpha^2(1-\xi)}{(1+\kappa)^2-\xi} > \delta. \tag{44}$$

It follows that for $\xi_{l_e} < \xi < 1$ and

$$\delta < \delta_{h_e}(\kappa, \xi) := \min\left\{\delta_h(\kappa, \xi), \delta_{h_{SD}}(\kappa), \frac{\alpha^2(1-\xi)}{(1+\kappa)^2 - \xi}\right\} \tag{45}$$

transition to FPP reduces efforts to create data set.

As δ_{0_1} does not depend on κ or ξ, $\delta_h(\kappa, \xi)$ depends negatively on both of them and $\delta_{h_{SD}}(\kappa)$ depends negatively on κ and does not depend on ξ, $\delta_{h_e}(\kappa, \xi)$ obviously depends nonpositively on κ and on ξ as well, since

$$\frac{\partial\left(\frac{\alpha^2(1-\xi)}{(1+\kappa)^2-\xi}\right)}{\partial\xi} = \frac{-\alpha^2[(1+\kappa)^2 - \xi] + \alpha^2(1-\xi)}{[(1+\kappa)^2 - \xi]^2} = \frac{-\alpha^2[(1+\kappa)^2 - 1]}{[(1+\kappa)^2 - \xi]^2} \leq 0. \tag{46}$$

(3) We want to show that transition to FPP reduces efforts to create data set if $\xi > \xi_h = \max\{\xi_{h_1}, \xi_{h_2}\}$ and $\delta > \max\{\delta_{h_1}, \delta_{h_{SD}}\}$. In this case, the Researcher chooses to delay strategically under FPP (since $\delta > \delta_{h_{SD}}$) or to share after 2 publications under NP.

To prove the last statement recall the arguments from the proof of Proposition 4(i). For $\xi > \xi_{h_1}$ it holds that $U_{R,NP_2^{share}} > U_{R,NP_2^{never}}$. For $\delta > \delta_{h_1}$ it holds that $U_{R,NP_2^{share}} > U_{R,NP_1^{share}}$ and for $\delta > \delta_{0_0}$ we have $U_{R,NP_1^{share}} > U_{R,NP_1^{never}}$. Further, $\delta_{h_1} > \delta_{0_0}$ whenever $\xi > \xi_{h_2}$ (see the arguments after (31)), so for $\xi > \xi_h$ and $\delta > \delta_{h_1}$ the Researcher chooses to share after 2 publications.

$e^*_{0,NP_2^{share}} > e^*_{0,FPP_{SD}}$ is equivalent to

$$\frac{\delta^3\gamma_D}{1 - \frac{\delta^2(1+\kappa)^2}{c^2\alpha^2} - \frac{\delta(1+\delta\kappa)^2}{\alpha^2}} > \frac{\delta^3\gamma_D}{1 - \frac{\delta^3(1+\kappa)^2}{\alpha^2} - \frac{\delta^2(1+\kappa)^2}{c^2\alpha^2}} \Leftrightarrow -\frac{\delta^3(1+\kappa)^2}{\alpha^2} > -\frac{\delta(1+\delta\kappa)^2}{\alpha^2} \Leftrightarrow$$
$$(1+\delta\kappa)^2 > \delta^2(1+\kappa)^2 \Leftrightarrow 1 + 2\delta\kappa + \delta^2\kappa^2 > \delta^2 + 2\delta^2\kappa + \delta^2\kappa^2 \Leftrightarrow$$
$$1 + 2\delta\kappa > \delta\left(\delta + 2\delta\kappa\right), \tag{47}$$

which is true for all $\delta \in (0, 1]$. So whenever the Researcher chooses to delay strategically under FPP and to share after 2 publications under NP, transition to FPP reduces efforts to create data set.

(ii) Transition to FPP does not influence efforts to create data set whenever the Researcher chooses to share after one publication under NP and not to delay strategically under FPP, i.e. whenever transition to FPP does not influence welfare (see Proposition 4(i)).

(iii)(1) As we know from the proof of Proposition 1 and from Proposition 3, for $\xi < 1$ and $\delta < \min\{\delta_h, \delta_{h_{SD}}\}$ the Researcher chooses to never share after one publication under

NP or not to delay strategically under FPP. From the definition of δ_{h_e} (see (45)) it follows that if $\delta_{h_e} < \delta < \min\{\delta_h, \delta_{h_{SD}}\}$, then transition to FPP increases efforts to create data set. We show that for $\kappa > 0$ there always exist such δs.

Consider $\xi < 1$ near to 1 and small δ. As was shown in the proof of Proposition 4(i) (see text after equation (29)), with these parameters the Researcher chooses to make one publication and then either to share or not to share, so that the relevant condition in $\min\{\delta_h, \delta_{h_{SD}}\}$ is $U_{R,NP_1^{never}} \geq U_{R,NP_1^{share}}$, i.e.

$$\delta_{0_0}(\xi) = \frac{\alpha^2(1-\xi^2)}{(1+\kappa)^2 - \xi^2}, \tag{48}$$

since δ_{0_1} and $\delta_{h_{SD}}$ are positive and do not depend on ξ, whereas $\delta_{0_0}(1) = 0$.

Notice that for $\xi = 1$ it holds that $\delta_{h_e} = 0$. So both $\delta_{0_0}(\xi)$ and $\delta_{h_e}(\xi)$ continuously approach 0 as $\xi \to 1$. In the neighborhood of $\xi = 1$ it holds that $\delta_{0_0}(\xi) > \delta_{h_e}(\xi)$, since

$$\begin{gathered}
\frac{\alpha^2(1-\xi^2)}{(1+\kappa)^2 - \xi^2} > \frac{\alpha^2(1-\xi)}{(1+\kappa)^2 - \xi} \;\Leftrightarrow \\
(1+\kappa)^2 - \xi - (1+\kappa)^2\xi^2 + \xi^3 > (1+\kappa)^2 - \xi^2 - (1+\kappa)^2\xi + \xi^3 \;\Leftrightarrow \\
\xi + (1+\kappa)^2 > 1 + (1+\kappa)^2\xi \;\Leftrightarrow (1+\kappa)^2(1-\xi) > 1 - \xi \;\Leftrightarrow \\
(1+\kappa)^2 > 1,
\end{gathered} \tag{49}$$

which is true for $\kappa > 0$.

For $\xi < \xi_{l_e}$ transition to FPP reduces efforts to create data set (see this Proposition, (i)(1)) and it follows that for $\xi_{l_e} < \xi < 1$ and $\delta_{h_e} < \delta < \min\{\delta_h, \delta_{h_{SD}}\}$ transition to FPP increases efforts to create data set.

(2) For $\xi < \xi_{h_1}$ (see (29)) and $\delta > \max\{\delta_{0_1}, \delta_{h_{SD}}, \delta_{h_1}\}$ the Researcher chooses to never share after two publications under NP or to delay strategically under FPP. Further we know from this proof, (i)(1), that in this case FPP increases efforts to create data set if and only if

$$\xi_{ul_e}(\delta) := \frac{1 - \frac{\delta^3(1+\kappa)^2}{\alpha^2} - \frac{\delta^2(1+\kappa)^2}{c^2\alpha^2}}{1 - \frac{\delta}{\alpha^2} - \frac{\delta^2}{c^2\alpha^2}} < \xi. \tag{50}$$

Since a square root of a number between 0 and 1 is bigger than the number itself, for $\delta = 1$ and $\kappa > 0$ it holds $\xi_{ul_e} < \xi_{h_1}$. It follows that FPP increases efforts to create a data set when $\xi_{ul_e}(\delta) < \xi < \xi_{h_1}$ and $\delta > \max\{\delta_{0_1}, \delta_{h_{SD}}, \delta_{h_1}\}$, whereby the set of such (ξ, δ) is always nonempty.

References

[1] Allan, R., 2012. Editorial: Geoscience data. Geoscience Data Journal. Retrieved from http://onlinelibrary.wiley.com/doi/10.1002/gdj.3/full (6 June 2014).

[2] Altman, M. and G. King (2007), A proposed standard for the scholarly citation of quantitative data, D-Lib Magazine, Vol. 13, No. 3/4.

[3] Anderson, R.G., W.H. Greene, B.D. McCullough and H.D. Vinod (2008), The role of data/code archives in the future of economic research, Journal of Economic Methodology, 15, 99-119.

[4] Andreoli Versbach, P. and F. Mueller-Langer, Open access to data: An ideal professed but not practised, Research Policy, forthcoming.

[5] Angrist, J.D. and J.-S. Pischke (2010), The credibility revolution in empirical economics: how better research design is taking the con out of econometrics, Journal of Economic Perspectives, 24(2), 3-30.

[6] Bernanke, B.S. (2004). Editorial statement, American Economic Review, 94, 404.

[7] Buckheit, J.B. and D.L. Donoho (1995), WaveLab and reproducible research in wavelets in statistics, eds.: A. Antoniadis and G. Oppenheim, New York: Springer-Verlag, 55-82.

[8] Coase, R. H. (1960), The problem of social cost, Journal of Law and Economics, 3, 1-44.

[9] Coupé, T., V. Smeets and F. Warzynski (2006), Incentives, sorting and productivity along the career: Evidence from a sample of top economists, Journal of Law, Economics, and Organization, 22(1), 137-167.

[10] Costello, M.J. (2009), Motivating online publication of data, BioScience, 59(5), 418-427.

[11] Dasgupta, P. and P.A. David (1994), Toward a new economics of science, Research Policy, 23(5), 487-521.

[12] Dewald, W.G., J. Thursby and R.G. Anderson (1986), Replication in empirical economics: The Journal of Money, Credit and Banking Project, American Economic Review, 76, 587-603.

[13] Economic and Social Research Council, 2010. ESRC research data policy. Retrieved from http://www.esrc.ac.uk/about-esrc/information/data-policy.aspx (5 March 2014).

[14] European Commission, 2012. Towards better access to scientific information: Boosting the benefits of public investments in research. European Commission: Brussels.

[15] Feigenbaum, S. and D.M. Levy (1993), The market for (ir)reproducible econometrics, Social Epistemology, 7(3), 215-232.

[16] Fienberg, S.E., M.E. Martin and M.L. Straf (1985), Sharing research data, National Academy Press: Washington, DC.

[17] Fudenberg, D., R. Gilbert, J. Stiglitz and J. Tirole (1983), Preemption, leapfrogging and competition in patent races, European Economic Review 22, 3-31.

[18] Gans, S.J. and S. Stern (2010), Is there a market for ideas?, Industrial and Corporate Change, 19 (3), 805-837.

[19] Glenn, D. (2008), Dispute over the economics of file sharing intensifies, The Chronicle of Higher Education, July 17, 2008, available at: http://chronicle.com/article/Dispute-Over-the-Economics-/989/.

[20] Haeussler, C. (2011), Information-sharing in academia and the industry: A comparative study, Research Policy, 40(1), 105-122.

[21] Haeussler, C., L. Jiang, J.G. Thursby and M. Thursby, 2014. Specific and general information sharing among competing academic researchers, Research Policy, 43(3), 465-475.

[22] Hall, R.E., 2009. Managing your career as an economist after tenure. CSWEP Newsletter, Winter 2009. 4-5.

[23] Hamermesh, D.S. (1997), Some thoughts on replications and reviews, Labour Economics, 4(2), 107-109.

[24] Kim, Y. and J.M. Stanton (2012), Institutional and individual influences on scientists' data sharing practices, Journal of Computational Science Education, 3(1), 47-56.

[25] Lacetera, N. and L. Zirulia (2011), The economics of scientific misconduct, Journal of Law, Economics, and Organization, 27(3), 568-603.

[26] McCullough, B.D. (2009), Open access economics journals and the market for reproducible economic research, Economic Analysis & Policy, 39(1), 118-126.

[27] McCullough, B.D. and H.D. Vinod (2003), Verifying the solution from a nonlinear solver: a case study, American Economic Review, 93, 873-892.

[28] McCullough, B.D., K.A. McGeary and T. Harrison (2006), Lessons from the JMCB Archive, Journal of Money, Credit and Banking, 38(4), 1093-1107.

[29] McCullough, B.D., K.A. McGeary and T.D. Harrison (2008), Do economics journal archives promote replicable research?, Canadian Journal of Economics, 41, 1406-1420.

[30] Mirowski, P. and S. Sklivas (1991), Why econometricians don't replicate (although they do reproduce), Review of Political Economy, 3(2), 146-163.

[31] Moffitt, R.A. (2007), Research data integrity in economics and other social sciences, Presentation to National Academies Committee on Science, Engineering, and Public Policy, Committee on Assuring the Integrity of Research Data, April, 2007.

[32] Moffitt, R.A. (2011), Report of the Editor: American Economic Review (with Appendix by P. J. Glandon), American Economic Review Papers and Proceedings, 101(3), 684-93.

[33] Mukherjee, A. and S. Stern (2009), Disclosure or secrecy? The dynamics of open science, International Journal of Industrial Organization, 27(3), 449-462.

[34] National Institutes of Health, 2003. NIH data sharing policy. Retrieved from http://grants1.nih.gov/grants/policy/data_sharing/ (5 March 2014).

[35] National Science Foundation, 2011. NSF grant proposal guide. Retrieved from http://www.nsf.gov/pubs/policydocs/pappguide/nsf11001/gpg_2.jsp#dmp (5 March 2014).

[36] Nature, 2013. Editorial announcement: Launch of an online data journal. Nature, 502(7470), 142.

[37] Nelson, B. (2009), Empty archives, Nature, 461, 160-163.

[38] Organisation for Economic Co-operation and Development, 2007. OECD principles and guidelines for access to research data from public funding. OECD: Paris.

[39] Perrino, T. et al. (2013), Advancing science through collaborative data sharing and synthesis, Perspectives on Psychological Science, 8, 433-444.

[40] Siegfried, J.J. and K.J White (1973), Financial rewards to research and teaching: A case study of academic economists, American Economic Review, 63(2), 309-315.

[41] Stephan, P.E. (1996), The economics of science, Journal of Economic Literature, 34(3), 1199-1235.

[42] Tuckman, H.P. and J. Leahey (1973), What is an article worth?, Journal of Political Economy, 83(5), 951-968.

[43] US House of Representatives, 2007. America Creating Opportunities to Meaningfully Promote Excellence in Technology, Education, and Science Act, 110th Congress, 1st Session. H.R. 2272, Section 1009, Release of scientific research results, Government Printing Office: Washington.

[44] Vlaeminck, S. and G.G. Wagner, 2014. On the role of research data centres in the management of publication-related research data, LIBER Quarterly, 23(4), 336-357.

[45] Wellcome Trust, 2007. Policy on data management and sharing. Retrieved from http://www.wellcome.ac.uk/About-us/Policy/Policy-and-position-statements/WTX035043.htm (5 March 2014).